

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(S1) International Patent Classification 7 : C07H 21/04, C07K 1/00, 14/00, C12N 1/21, 15/00, 15/09, 15/63, 15/70, C12P 19/34		A1	(11) International Publication Number: WO 00/52027 (43) International Publication Date: 8 September 2000 (08.09.00)
(21) International Application Number: PCT/US00/05432 (22) International Filing Date: 2 March 2000 (02.03.00) (30) Priority Data: 60/122,389 2 March 1999 (02.03.99) US 60/126,049 23 March 1999 (23.03.99) US 60/136,744 28 May 1999 (28.05.99) US		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(71) Applicant: LIFE TECHNOLOGIES, INC. [US/US]; 9800 Medical Center Drive, Rockville, MD 20850 (US). (72) Inventors: HARTLEY, James, L.; 7409 Hillside Drive, Frederick, MD 21702 (US). BRASCH, Michael, A.; 20931 Sunnacres Road, Gaithersburg, MD 20882 (US). TEMPLE, Gary, F.; 114 Ridge Road, Washington Grove, MD 20882 (US). CHEO, David; 2006 Baltimore Road, #21, Rockville, MD 20851 (US). (74) Agents: ESMOND, Robert, W. et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100 New York Avenue, N.W., Washington, DC 20005-3934 (US).		Published <i>With international search report.</i> <i>With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.</i>	
(54) Title: COMPOSITIONS AND METHODS FOR USE IN RECOMBINATIONAL CLONING OF NUCLEIC ACIDS			
(57) Abstract			
<p>The present invention relates generally to compositions and methods for use in recombinational cloning of nucleic acid molecules. In particular, the invention relates to nucleic acid molecules encoding one or more recombination sites or portions thereof, to nucleic acid molecules comprising one or more of these recombination site nucleotide sequences and optionally comprising one or more additional physical or functional nucleotide sequences. The invention also relates to vectors comprising the nucleic acid molecules of the invention, to host cells comprising the vectors or nucleic acid molecules of the invention, to methods of producing polypeptides using the nucleic acid molecules of the invention, and to polypeptides encoded by these nucleic acid molecules or produced by the methods of the invention. The invention also relates to antibodies that bind to one or more polypeptides of the invention or epitopes thereof. The invention also relates to the use of these compositions in methods for recombinational cloning of nucleic acids, <i>in vitro</i> and <i>in vivo</i>, to provide chimeric DNA molecules that have particular characteristics and/or DNA segments.</p>			

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LJ	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

Compositions and Methods for Use in Recombinational Cloning of Nucleic Acids

5

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates generally to recombinant DNA technology.
10 More particularly, the present invention relates to compositions and methods for use in recombinational cloning of nucleic acid molecules. The invention relates specifically to nucleic acid molecules encoding one or more recombination sites or one or more partial recombination sites, particularly *attB*, *attP*, *attL*, and *attR*, and fragments, mutants, variants and derivatives thereof. The invention also relates to such nucleic acid molecules wherein the one or more recombination site nucleotide sequences is operably linked to the one or more additional physical or functional nucleotide sequences. The invention also relates to vectors comprising the nucleic acid molecules of the invention, to host cells comprising the vectors or nucleic acid molecules of the invention, to methods of producing polypeptides and RNAs encoded by the nucleic acid molecules of the invention, and to polypeptides encoded by these nucleic acid molecules or produced by the methods of the invention, which may be fusion proteins. The invention also relates to antibodies that bind to one or more polypeptides of the invention or epitopes thereof, which may be monoclonal or polyclonal antibodies. The invention also relates to the use of these nucleic acid molecules, vectors, polypeptides and antibodies in methods for recombinational cloning of nucleic acids, *in vitro* and *in vivo*, to provide chimeric DNA molecules that have particular characteristics and/or DNA segments. More particularly, the antibodies of the invention may be used to identify and/or purify proteins or fusion proteins encoded by the nucleic acid molecules or vectors of the invention, or to identify and/or purify the nucleic acid molecules of the invention.

15

20

25

30

Related Art

5 **Site-specific recombinases.** Site-specific recombinases are proteins that are present in many organisms (e.g. viruses and bacteria) and have been characterized to have both endonuclease and ligase properties. These recombinases (along with associated proteins in some cases) recognize specific sequences of bases in DNA and exchange the DNA segments flanking those segments. The recombinases and associated proteins are collectively referred to as "recombination proteins" (see, e.g., Landy, A., *Current Opinion in Biotechnology* 3:699-707 (1993)).

10

Numerous recombination systems from various organisms have been described. See, e.g., Hoess *et al.*, *Nucleic Acids Research* 14(6):2287 (1986); Abremski *et al.*, *J. Biol. Chem.* 261(1):391 (1986); Campbell, J. *Bacteriol.* 174(23):7495 (1992); Qian *et al.*, *J. Biol. Chem.* 267(11):7794 (1992);
15 Araki *et al.*, *J. Mol. Biol.* 225(1):25 (1992); Maeser and Kahnmann *Mol. Gen. Genet.* 230:170-176 (1991); Esposito *et al.*, *Nucl. Acids Res.* 25(18):3605 (1997).

20 Many of these belong to the integrase family of recombinases (Argos *et al.* *EMBO J.* 5:433-440 (1986); Vozianov *et al.*, *Nucl. Acids Res.* 27:930 (1999)).

Perhaps the best studied of these are the Integrase/att system from bacteriophage λ (Landy, A. *Current Opinions in Genetics and Devel.* 3:699-707 (1993)), the Cre/loxP system from bacteriophage P1 (Hoess and Abremski (1990) In *Nucleic Acids and Molecular Biology*, vol. 4. Eds.: Eckstein and Lilley, Berlin-Heidelberg: Springer-Verlag; pp. 90-109), and the FLP/FRT system from the Saccharomyces cerevisiae 2 μ circle plasmid (Broach *et al.* *Cell* 29:227-234 (1982)).

25

Backman (U.S. Patent No. 4,673,640) discloses the *in vivo* use of λ recombinase to recombine a protein producing DNA segment by enzymatic site-specific recombination using wild-type recombination sites attB and attP.

30 Hasan and Szybalski (*Gene* 56:145-151 (1987)) discloses the use of λ Int recombinase *in vivo* for intramolecular recombination between wild type attP and attB sites which flank a promoter. Because the orientations of these sites are

inverted relative to each other, this causes an irreversible flipping of the promoter region relative to the gene of interest.

5 Palazzolo *et al.* *Gene* 88:25-36 (1990), discloses phage lambda vectors having bacteriophage λ arms that contain restriction sites positioned outside a cloned DNA sequence and between wild-type *loxP* sites. Infection of *E. coli* cells that express the Cre recombinase with these phage vectors results in recombination between the *loxP* sites and the *in vivo* excision of the plasmid replicon, including the cloned cDNA.

10 Pósfai *et al.* (*Nucl. Acids Res.* 22:2392-2398 (1994)) discloses a method for inserting into genomic DNA partial expression vectors having a selectable marker, flanked by two wild-type FRT recognition sequences. FLP site-specific recombinase as present in the cells is used to integrate the vectors into the genome at predetermined sites. Under conditions where the replicon is functional, this cloned genomic DNA can be amplified.

15 Bebee *et al.* (U.S. Patent No. 5,434,066) discloses the use of site-specific recombinases such as Cre for DNA containing two *loxP* sites for *in vivo* recombination between the sites.

20 Boyd (*Nucl. Acids Res.* 21:817-821 (1993)) discloses a method to facilitate the cloning of blunt-ended DNA using conditions that encourage intermolecular ligation to a dephosphorylated vector that contains a wild-type *loxP* site acted upon by a Cre site-specific recombinase present in *E. coli* host cells.

25 Waterhouse *et al.* (WO 93/19172 and *Nucleic Acids Res.* 21 (9):2265 (1993)) disclose an *in vivo* method where light and heavy chains of a particular antibody were cloned in different phage vectors between *loxP* and *loxP 511* sites and used to transfet new *E. coli* cells. Cre, acting in the host cells on the two parental molecules (one plasmid, one phage), produced four products in equilibrium: two different cointegrates (produced by recombination at either *loxP* or *loxP 511* sites), and two daughter molecules, one of which was the desired product.

30 Schlake & Bode (*Biochemistry* 33:12746-12751 (1994)) discloses an *in vivo* method to exchange expression cassettes at defined chromosomal locations, each flanked by a wild type and a spacer-mutated FRT recombination site. A

-4-

double-reciprocal crossover was mediated in cultured mammalian cells by using this FLP/FRT system for site-specific recombination.

5 Hartley *et al.* (U.S. Patent No. 5,888,732) disclose compositions and methods for recombinational exchange of nucleic acid segments and molecules, including for use in recombinational cloning of a variety of nucleic acid molecules *in vitro* and *in vivo*, using a combination of wildtype and mutated recombination sites and recombination proteins.

10 *Transposases.* The family of enzymes, the transposases, has also been used to transfer genetic information between replicons. Transposons are structurally variable, being described as simple or compound, but typically encode the recombinase gene flanked by DNA sequences organized in inverted orientations. Integration of transposons can be random or highly specific. Representatives such as Tn7, which are highly site-specific, have been applied to the *in vivo* movement of DNA segments between replicons (Lucklow *et al.*,
15 *J. Virol.* 67:4566-4579 (1993)).

20 Devine and Boeke *Nucl. Acids Res.* 22:3765-3772 (1994), discloses the construction of artificial transposons for the insertion of DNA segments, *in vitro*, into recipient DNA molecules. The system makes use of the integrase of yeast TY1 virus-like particles. The DNA segment of interest is cloned, using standard methods, between the ends of the transposon-like element TY1. In the presence
25 of the TY1 integrase, the resulting element integrates randomly into a second target DNA molecule.

30 *Recombination Sites.* Also key to the integration/recombination reactions mediated by the above-noted recombination proteins and/or transposases are recognition sequences, often termed "recombination sites," on the DNA molecules participating in the integration/recombination reactions. These recombination sites are discrete sections or segments of DNA on the participating nucleic acid molecules that are recognized and bound by the recombination proteins during the initial stages of integration or recombination. For example, the recombination site for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See Figure 1 of Sauer, B., *Curr. Opin. Biotech.*

5 5:521-527 (1994). Other examples of recognition sequences include the *attB*, *attP*, *attL*, and *attR* sequences which are recognized by the recombination protein λ Int. *attB* is an approximately 25 base pair sequence containing two 9 base pair core-type Int binding sites and a 7 base pair overlap region, while *attP* is an approximately 240 base pair sequence containing core-type Int binding sites and arm-type Int binding sites as well as sites for auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). See Landy, *Curr. Opin. Biotech.* 3:699-707 (1993); see also U.S. Patent No. 5,888,732, which is incorporated by reference herein.

10 15 **DNA cloning.** The cloning of DNA segments currently occurs as a daily routine in many research labs and as a prerequisite step in many genetic analyses. The purpose of these clonings is various, however, two general purposes can be considered: (1) the initial cloning of DNA from large DNA or RNA segments (chromosomes, YACs, PCR fragments, mRNA, etc.), done in a relative handful of known vectors such as pUC, pGem, pBlueScript, and (2) the subcloning of these DNA segments into specialized vectors for functional analysis. A great deal of time and effort is expended both in the transfer of DNA segments from the initial cloning vectors to the more specialized vectors. This transfer is called subcloning.

20 The basic methods for cloning have been known for many years and have changed little during that time. A typical cloning protocol is as follows:

- (1) digest the DNA of interest with one or two restriction enzymes;
- (2) gel purify the DNA segment of interest when known;
- (3) prepare the vector by cutting with appropriate restriction enzymes, treating with alkaline phosphatase, gel purify etc., as appropriate;
- (4) ligate the DNA segment to the vector, with appropriate controls to eliminate background of uncut and self-ligated vector;
- (5) introduce the resulting vector into an *E. coli* host cell;
- (6) pick selected colonies and grow small cultures overnight;
- (7) make DNA minipreps; and

-6-

(8) analyze the isolated plasmid on agarose gels (often after diagnostic restriction enzyme digestions) or by PCR.

The specialized vectors used for subcloning DNA segments are functionally diverse. These include but are not limited to: vectors for expressing nucleic acid molecules in various organisms; for regulating nucleic acid molecule expression; for providing tags to aid in protein purification or to allow tracking of proteins in cells; for modifying the cloned DNA segment (e.g., generating deletions); for the synthesis of probes (e.g., riboprobes); for the preparation of templates for DNA sequencing; for the identification of protein coding regions; for the fusion of various protein-coding regions; to provide large amounts of the DNA of interest, etc. It is common that a particular investigation will involve subcloning the DNA segment of interest into several different specialized vectors.

As known in the art, simple subclonings can be done in one day (e.g., the DNA segment is not large and the restriction sites are compatible with those of the subcloning vector). However, many other subclonings can take several weeks, especially those involving unknown sequences, long fragments, toxic genes, unsuitable placement of restriction sites, high backgrounds, impure enzymes, etc. Subcloning DNA fragments is thus often viewed as a chore to be done as few times as possible.

Several methods for facilitating the cloning of DNA segments have been described, e.g., as in the following references.

Ferguson, J., et al. *Gene* 16:191 (1981), discloses a family of vectors for subcloning fragments of yeast DNA. The vectors encode kanamycin resistance. Clones of longer yeast DNA segments can be partially digested and ligated into the subcloning vectors. If the original cloning vector conveys resistance to ampicillin, no purification is necessary prior to transformation, since the selection will be for kanamycin.

Hashimoto-Gotoh, T., et al. *Gene* 41:125 (1986), discloses a subcloning vector with unique cloning sites within a streptomycin sensitivity gene; in a streptomycin-resistant host, only plasmids with inserts or deletions in the dominant sensitivity gene will survive streptomycin selection.

5

10

15

20

25

Accordingly, traditional subcloning methods, using restriction enzymes and ligase, are time consuming and relatively unreliable. Considerable labor is expended, and if two or more days later the desired subclone can not be found among the candidate plasmids, the entire process must then be repeated with alternative conditions attempted. Although site specific recombinases have been used to recombine DNA *in vivo*, the successful use of such enzymes *in vitro* was expected to suffer from several problems. For example, the site specificities and efficiencies were expected to differ *in vitro*; topologically linked products were expected; and the topology of the DNA substrates and recombination proteins was expected to differ significantly *in vitro* (see, e.g., Adams *et al*, *J. Mol. Biol.* 226:661-73 (1992)). Reactions that could go on for many hours *in vivo* were expected to occur in significantly less time *in vitro* before the enzymes became inactive. In addition, the stabilities of the recombination enzymes after incubation for extended periods of time in *in vitro* reactions was unknown, as were the effects of the topologies (*i.e.*, linear, coiled, supercoiled, etc.) of the nucleic acid molecules involved in the reaction. Multiple DNA recombination products were expected in the biological host used, resulting in unsatisfactory reliability, specificity or efficiency of subcloning. Thus, *in vitro* recombination reactions were not expected to be sufficiently efficient to yield the desired levels of product.

Accordingly, there is a long felt need to provide an alternative subcloning system that provides advantages over the known use of restriction enzymes and ligases.

30

SUMMARY OF THE INVENTION

The present invention relates to nucleic acid molecules encoding one or more recombination sites or one or more partial recombination sites, particularly *attB*, *attP*, *attL*, and *attR*, and fragments, mutants, variants and derivatives thereof. The invention also relates to such nucleic acid molecules comprising one or more of the recombination site nucleotide sequences or portions thereof and one or more additional physical or functional nucleotide sequences, such as those

5

10

15

20

25

30

encoding one or more multiple cloning sites, one or more transcription termination sites, one or more transcriptional regulatory sequences (e.g., one or more promoters, enhancers, or repressors), one or more translational signal sequences, one or more nucleotide sequences encoding a fusion partner protein or peptide (e.g., GST, His₆ or thioredoxin), one or more selection markers or modules, one or more nucleotide sequences encoding localization signals such as nuclear localization signals or secretion signals, one or more origins of replication, one or more protease cleavage sites, one or more desired proteins or peptides encoded by a gene or a portion of a gene, and one or more 5' or 3' polynucleotide tails (particularly a poly-G tail). The invention also relates to such nucleic acid molecules wherein the one or more recombination site nucleotide sequences is operably linked to the one or more additional physical or functional nucleotide sequences.

The invention also relates to primer nucleic acid molecules comprising the recombination site nucleotide sequences of the invention (or portions thereof), and to such primer nucleic acid molecules linked to one or more target-specific (e.g., one or more gene-specific) primer nucleic acid sequences. Such primers may also comprise sequences complementary or homologous to DNA or RNA sequences to be amplified, e.g., by PCR, RT-PCR, etc. Such primers may also comprise sequences or portions of sequences useful in the expression of protein genes (ribosome binding sites, localization signals, protease cleavage sites, repressor binding sites, promoters, transcription stops, stop codons, etc.). Said primers may also comprise sequences or portions of sequences useful in the manipulation of DNA molecules (restriction sites, transposition sites, sequencing primers, etc.). The primers of the invention may be used in nucleic acid synthesis and preferably are used for amplification (e.g., PCR) of nucleic acid molecules. When the primers of the invention include target- or gene-specific sequences (any sequence contained within the target to be synthesized or amplified including translation signals, gene sequences, stop codons, transcriptional signals (e.g., promoters) and the like), amplification or synthesis of target sequences or genes may be accomplished. Thus, the invention relates to synthesis of a nucleic acid molecules comprising mixing one or more primers of the invention with a nucleic acid

template, and incubating said mixture under conditions sufficient to make a first nucleic acid molecule complementary to all or a portion of said template. Thus, the invention relates specifically to a method of synthesizing a nucleic acid molecule comprising:

- 5 (a) mixing a nucleic acid template with a polypeptide having polymerase activity and one or more primers comprising one or more recombination sites or portions thereof; and
- 10 (b) incubating said mixture under conditions sufficient to synthesize a first nucleic acid molecule complementary to all or a portion of said template and which preferably comprises one or more recombination sites or portions thereof.

15 Such method of the invention may further comprise incubating said first synthesized nucleic acid molecule under conditions sufficient to synthesize a second nucleic acid molecule complementary to all or a portion of said first nucleic acid molecule. Such synthesis may provide for a first nucleic acid molecule having a recombination site or portion thereof at one or both of its termini.

20 In a preferred aspect, for the synthesis of the nucleic acid molecules, at least two primers are used wherein each primer comprises a homologous sequence at its terminus and/or within internal sequences of each primer (which may have a homology length of about 2 to about 500 bases, preferably about 3 to about 100 bases, about 4 to about 50 bases, about 5 to about 25 bases and most preferably about 6 to about 18 base overlap). In a preferred aspect, the first such primer comprises at least one target-specific sequence and at least one recombination site or portion thereof while the second primer comprises at least one recombination site or portion thereof. Preferably, the homologous regions between the first and second primers comprise at least a portion of the recombination site. In another aspect, the homologous regions between the first and second primers may comprise one or more additional sequences, e.g., expression signals, translational start motifs, or other sequences adding functionality to the desired nucleic acid sequence upon amplification. In practice, two pairs of primers prime synthesis or amplification of a nucleic acid molecule. In a preferred aspect, all or at least a portion of the synthesized or amplified nucleic acid molecule will be homologous

-10-

to all or a portion of the template and further comprises a recombination site or a portion thereof at at least one terminus and preferably both termini of the synthesized or amplified molecule. Such synthesized or amplified nucleic acid molecule may be double stranded or single stranded and may be used in the recombinational cloning methods of the invention. The homologous primers of the invention provide a substantial advantage in that one set of the primers may be standardized for any synthesis or amplification reaction. That is, the primers providing the recombination site sequences (without the target specific sequences) can be pre-made and readily available for use. This in practice allows the use of shorter custom made primers that contain the target specific sequence needed to synthesize or amplify the desired nucleic acid molecule. Thus, this provides reduced time and cost in preparing target specific primers (e.g., shorter primers containing the target specific sequences can be prepared and used in synthesis reactions). The standardized primers, on the other hand, may be produced in mass to reduce cost and can be readily provided (e.g., in kits or as a product) to facilitate synthesis of the desired nucleic acid molecules.

Thus, in one preferred aspect, the invention relates to a method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template specific sequence (complementary to or capable of hybridizing to said templates) and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said second primer is homologous to or complementary to at least a portion of said first primer; and
- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one and preferably both termini of said molecules.

More specifically, the invention relates to a method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template specific sequence (complementary to or capable of hybridizing to said templates) and at least a portion of a recombination site, and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said recombination site on said second primer is complementary to or homologous to at least a portion of said recombination site on said first primer; and
- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one and preferably both termini of said molecules.

In a more preferred aspect, the invention relates to a method of amplifying or synthesizing one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and one or more first primers comprising at least a portion of a recombination site and a template specific sequence (complementary to or capable of hybridizing to said template);
- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more first nucleic acid molecules complementary to all or a portion of said templates wherein said molecules comprise at least a portion of a recombination site at one and preferably both termini of said molecules;
- (c) mixing said molecules with one or more second primers comprising one or more recombination sites, wherein said recombination sites of said second primers are homologous to or

-12-

complementary to at least a portion of said recombination sites on
said first nucleic acid molecules; and

- 5 (d) incubating said mixture under conditions sufficient to synthesize or
amplify one or more second nucleic acid molecules complementary
to all or a portion of said first nucleic acid molecules and which
comprise one or more recombination sites at one and preferably
both termini of said molecules.

10 The invention also relates to vectors comprising the nucleic acid molecules
of the invention, to host cells comprising the vectors or nucleic acid molecules of
the invention, to methods of producing polypeptides encoded by the nucleic acid
molecules of the invention, and to polypeptides encoded by these nucleic acid
molecules or produced by the methods of the invention, which may be fusion
proteins. The invention also relates to antibodies that bind to one or more
15 polypeptides of the invention or epitopes thereof, which may be monoclonal or
polyclonal antibodies. The invention also relates to the use of these nucleic acid
molecules, primers, vectors, polypeptides and antibodies in methods for
recombinational cloning of nucleic acids, *in vitro* and *in vivo*, to provide chimeric
DNA molecules that have particular characteristics and/or DNA segments.

20 The antibodies of the invention may have particular use to identify and/or
purify peptides or proteins (including fusion proteins produced by the invention),
and to identify and/or purify the nucleic acid molecules of the invention or
portions thereof.

25 The methods for *in vitro* or *in vivo* recombinational cloning of nucleic acid
molecule generally relate to recombination between at least a first nucleic acid
molecule having at least one recombination site and a second nucleic acid
molecule having at least one recombination site to provide a chimeric nucleic acid
molecule. In one aspect, the methods relate to recombination between a first
vector having at least one recombination site and a second vector having at least
one recombination site to provide a chimeric vector. In another aspect, a nucleic
30 acid molecule having at least one recombination site is combined with a vector
having at least one recombination site to provide a chimeric vector. In a most
preferred aspect, the nucleic acid molecules or vectors used in recombination

comprise two or more recombination sites. In a more specific embodiment of the invention, the recombination methods relate to a Destination Reaction (also referred to herein as an "LR reaction") in which recombination occurs between an Entry clone and a Destination Vector. Such a reaction transfers the nucleic acid molecule of interest from the Entry Clone into the Destination Vector to create an Expression Clone. The methods of the invention also specifically relate to an Entry or Gateward reaction (also referred to herein as a "BP reaction") in which an Expression Clone is recombined with a Donor vector to produce an Entry clone. In other aspects, the invention relates to methods to prepare Entry clones by combining an Entry vector with at least one nucleic acid molecule (e.g., gene or portion of a gene). The invention also relates to conversion of a desired vector into a Destination Vector by including one or more (preferably at least two) recombination sites in the vector of interest. In a more preferred aspect, a nucleic acid molecule (e.g., a cassette) having at least two recombination sites flanking a selectable marker (e.g., a toxic gene or a genetic element preventing the survival of a host cell containing that gene or element, and/or preventing replication, partition or heritability of a nucleic acid molecule (e.g., a vector or plasmid) comprising that gene or element) is added to the vector to make a Destination Vector of the invention.

Preferred vectors for use in the invention include prokaryotic vectors, eukaryotic vectors, or vectors which may shuttle between various prokaryotic and/or eukaryotic systems (e.g. shuttle vectors). Preferred prokaryotic vectors for use in the invention include but are not limited to vectors which may propagate and/or replicate in gram negative and/or gram positive bacteria, including bacteria of the genera *Escherichia*, *Salmonella*, *Proteus*, *Clostridium*, *Klebsiella*, *Bacillus*, *Streptomyces*, and *Pseudomonas* and preferably in the species *E. coli*. Eukaryotic vectors for use in the invention include vectors which propagate and/or replicate in yeast cells, plant cells, mammalian cells, (particularly human and mouse), fungal cells, insect cells, nematode cells, fish cells and the like. Particular vectors of interest include but are not limited to cloning vectors, sequencing vectors, expression vectors, fusion vectors, two-hybrid vectors, gene therapy vectors, phage display vectors, gene-targeting vectors, PACs, BACs, YACs, MACs, and

reverse two-hybrid vectors. Such vectors may be used in prokaryotic and/or eukaryotic systems depending on the particular vector.

In another aspect, the invention relates to kits which may be used in carrying out the methods of the invention, and more specifically relates to cloning or subcloning kits and kits for carrying out the LR Reaction (e.g., making an Expression Clone), for carrying out the BP Reaction (e.g., making an Entry Clone), and for making Entry Clone and Destination Vector molecules of the invention. Such kits may comprise a carrier or receptacle being compartmentalized to receive and hold therein any number of containers. Such containers may contain any number of components for carrying out the methods of the invention or combinations of such components. In particular, a kit of the invention may comprise one or more components (or combinations thereof) selected from the group consisting of one or more recombination proteins or auxiliary factors or combinations thereof, one or more compositions comprising one or more recombination proteins or auxiliary factors or combinations thereof (for example, GATEWAY™ LR Clonase™ Enzyme Mix or GATEWAY™ BP Clonase™ Enzyme Mix), one or more reaction buffers, one or more nucleotides, one or more primers of the invention, one or more restriction enzymes, one or more ligases, one or more polypeptides having polymerase activity (e.g., one or more reverse transcriptases or DNA polymerases), one or more proteinases (e.g., proteinase K or other proteinases), one or more Destination Vector molecules, one or more Entry Clone molecules, one or more host cells (e.g. competent cells, such as *E. coli* cells, yeast cells, animal cells (including mammalian cells, insect cells, nematode cells, avian cells, fish cells, etc.), plant cells, and most particularly *E. coli* DB3.1 host cells, such as *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells), instructions for using the kits of the invention (e.g., to carry out the methods of the invention), and the like. In related aspects, the kits of the invention may comprise one or more nucleic acid molecules encoding one or more recombination sites or portions thereof, particularly one or more nucleic acid molecules comprising a nucleotide sequence encoding the one or more recombination sites or portions thereof of the invention. Preferably, such nucleic acid molecules comprise at least two recombination sites which flank a selectable

marker (e.g., a toxic gene and/or antibiotic resistance gene). In a preferred aspect, such nucleic acid molecules are in the form of a cassette (e.g., a linear nucleic acid molecule comprising one or more and preferably two or more recombination sites or portions thereof).

5 Kits for inserting or adding recombination sites to nucleic acid molecules of interest may comprise one or more nucleases (preferably restriction endonucleases), one or more ligases, one or more topoisomerases, one or more polymerases, and one or more nucleic acid molecules or adapters comprising one or more recombination sites. Kits for integrating recombination sites into one or 10 more nucleic acid molecules of interest may comprise one or more components (or combinations thereof) selected from the group consisting of one or more integration sequences comprising one or more recombination sites. Such integration sequences may comprise one or more transposons, integrating viruses, homologous recombination sequences, RNA molecules, one or more host cells 15 and the like.

Kits for making the Entry Clone molecules of the invention may comprise any or a number of components and the composition of such kits may vary depending on the specific method involved. Such methods may involve inserting the nucleic acid molecules of interest into an Entry or Donor Vector by the recombinational cloning methods of the invention, or using conventional molecular biology techniques (e.g., restriction enzyme digestion and ligation). In a preferred aspect, the Entry Clone is made using nucleic acid amplification or synthesis products. Kits for synthesizing Entry Clone molecules from amplification or synthesis products may comprise one or more components (or combinations thereof) selected from the group consisting of one or more Donor Vectors (e.g., 20 one or more attP vectors including, but not limited to, pDONR201 (Figure 49), pDONR202 (Figure 50), pDONR203 (Figure 51), pDONR204 (Figure 52), pDONR205 (Figure 53), pDONR206 (Figure 53), and the like), one or more polypeptides having polymerase activity (preferably DNA polymerases and most 25 preferably thermostable DNA polymerases), one or more proteinases, one or more reaction buffers, one or more nucleotides, one or more primers comprising one or 30

-16-

more recombination sites or portions thereof, and instructions for making one or more Entry Clones.

Kits for making the Destination vectors of the invention may comprise any number of components and the compositions of such kits may vary depending on the specific method involved. Such methods may include the recombination methods of the invention or conventional molecular biology techniques (e.g., restriction endonuclease digestion and ligation). In a preferred aspect, the Destination vector is made by inserting a nucleic acid molecule comprising at least one recombination site (or portion thereof) of the invention (preferably a nucleic acid molecule comprising at least two recombination sites or portions thereof flanking a selectable marker) into a desired vector to convert the desired vector into a Destination vector of the invention. Such kits may comprise at least one component (or combinations thereof) selected from the group consisting of one or more restriction endonucleases, one or more ligases, one or more polymerases, one or more nucleotides, reaction buffers, one or more nucleic acid molecules comprising at least one recombination site or portion thereof (preferably at least one nucleic acid molecule comprising at least two recombination sites flanking at least one selectable marker, such as a cassette comprising at least one selectable marker such as antibiotic resistance genes and/or toxic genes), and instructions for making such Destination vectors.

The invention also relates to kits for using the antibodies of the invention in identification and/or isolation of peptides and proteins (which may be fusion proteins) produced by the nucleic acid molecules of the invention, and for identification and/or isolation of the nucleic acid molecules of the invention or portions thereof. Such kits may comprise one or more components (or combination thereof) selected from the group consisting of one or more antibodies of the invention, one or more detectable labels, one or more solid supports and the like.

Other preferred embodiments of the present invention will be apparent to one of ordinary skill in light of what is known in the art, in light of the following drawings and description of the invention, and in light of the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

5 **Figure 1** depicts one general method of the present invention, wherein the starting (parent) DNA molecules can be circular or linear. The goal is to exchange the new subcloning vector D for the original cloning vector B. It is desirable in one embodiment to select for AD and against all the other molecules, including the Cointegrate. The square and circle are sites of recombination: *e.g.*, *lox* (such as *loxP*) sites, *att* sites, *etc.* For example, segment D can contain expression signals, protein fusion domains, new drug markers, new origins of replication, or specialized functions for mapping or sequencing DNA. It should be noted that the cointegrate molecule contains Segment D (Destination vector) adjacent to segment A (Insert), thereby juxtaposing functional elements in D with the insert in A. Such molecules can be used directly *in vitro* (*e.g.*, if a promoter is positioned adjacent to a gene-for *in vitro* transcription/translation) or *in vivo* (following isolation in a cell capable of propagating *ccdB*-containing vectors) by selecting for the selection markers in Segments B+D. As one skilled in the art will recognize, this single step method has utility in certain envisioned applications of the invention.

10

15 **Figure 2** is a more detailed depiction of the recombinational cloning system of the invention, referred to herein as the "GATEWAY™ Cloning System." This figure depicts the production of Expression Clones via a "Destination Reaction," which may also be referred to herein as an "LR Reaction." A *kan*^r vector (referred to herein as an "Entry clone") containing a DNA molecule of interest (*e.g.*, a gene) localized between an *attL1* site and an *attL2* site is reacted with an *amp*^r vector (referred to herein as a "Destination Vector") containing a toxic or "death" gene localized between an *attR1* site and an *attR2* site, in the presence of GATEWAY™ LR Clonase™ Enzyme Mix (a mixture of Int, IHF and Xis). After incubation at 25 °C for about 60 minutes, the reaction yields an *amp*^r Expression Clone containing the DNA molecule of interest localized between an *attB1* site and an *attB2* site, and a *kan*^r byproduct molecule, as well as intermediates. The reaction mixture may then be transformed into host cells (*e.g.*, *E. coli*) and clones containing the nucleic acid molecule of interest may

20

25

30

be selected by plating the cells onto ampicillin-containing media and picking amp^r colonies.

5

Figure 3 is a schematic depiction of the cloning of a nucleic acid molecule from an Entry clone into multiple types of Destination vectors, to produce a variety of Expression Clones. Recombination between a given Entry clone and different types of Destination vectors (not shown), via the LR Reaction depicted in Figure 2, produces multiple different Expression Clones for use in a variety of applications and host cell types.

10

15

20

25

Figure 4 is a detailed depiction of the production of Entry Clones via a "BP reaction," also referred to herein as an "Entry Reaction" or a "Gateward Reaction." In the example shown in this figure, an amp^r expression vector containing a DNA molecule of interest (e.g., a gene) localized between an attB1 site and an attB2 site is reacted with a kan^r Donor vector (e.g., an attP vector; here, GATEWAY™ pDONR201 (see Figure 49A-C)) containing a toxic or "death" gene localized between an attP1 site and an attP2 site, in the presence of GATEWAY™ BP Clonase™ Enzyme Mix (a mixture of Int and IHF). After incubation at 25°C for about 60 minutes, the reaction yields a kan^r Entry clone containing the DNA molecule of interest localized between an attL1 site and an attL2 site, and an amp^r by-product molecule. The Entry clone may then be transformed into host cells (e.g., *E. coli*) and clones containing the Entry clone (and therefore the nucleic acid molecule of interest) may be selected by plating the cells onto kanamycin-containing media and picking kan^r colonies. Although this figure shows an example of use of a kan^r Donor vector, it is also possible to use Donor vectors containing other selection markers, such as the gentamycin resistance or tetracycline resistance markers, as discussed herein.

Figure 5 is a more detailed schematic depiction of the LR ("Destination") reaction (Figure 5A) and the BP ("Entry" or "Gateward") reaction (Figure 5B) of the GATEWAY™ Cloning System, showing the reactants, products and byproducts of each reaction.

Figure 6 shows the sequences of the attB1 and attB2 sites flanking a gene of interest after subcloning into a Destination Vector to create an Expression Clone.

Figure 7 is a schematic depiction of four ways to make Entry Clones using the compositions and methods of the invention: 1. using restriction enzymes and ligase; 2. starting with a cDNA library prepared in an attL Entry Vector; 3. using an Expression Clone from a library prepared in an attB Expression Vector via the BxP reaction; and 4. recombinational cloning of PCR fragments with terminal attB sites, via the BxP reaction. Approaches 3 and 4 rely on recombination with a Donor vector (here, an attP vector such as pDONR201 (see Figure 49A-C), pDONR202 (see Figure 50A-C), pDONR203 (see Figure 51A-C), pDONR204 (see Figure 52A-C), pDONR205 (see Figure 53A-C), or pDONR206 (see Figure 54A-C), for example) that provides an Entry Clone carrying a selection marker such as kan^r, gen^r, tet^r, or the like.

Figure 8 is a schematic depiction of cloning of a PCR product by a BxP (Entry or Gateward) reaction. A PCR product with 25 bp terminal attB sites (plus four Gs) is shown as a substrate for the BxP reaction. Recombination between the attB-PCR product of a gene and a Donor vector (which donates an Entry Vector that carries kan^r) results in an Entry Clone of the PCR product.

Figure 9 is a listing of the nucleotide sequences of the recombination sites designated herein as attB1, attB2, attP1, attP2, attL1, attL2, attR1 and attR2. Sequences are written conventionally, from 5' to 3'.

Figures 10-20: The plasmid backbone for all the Entry Vectors depicted herein is the same, and is shown in Figure 10A for the Entry Vector pENTR1A. For other Entry Vectors shown in Figures 11-20, only the sequences shown in Figure "A" for each figure set (*i.e.*, Figure 11A, Figure 12A, etc.) are different (within the attL1-attL2 cassettes) from those shown in Figure 10 -- the plasmid backbone is identical.

Figure 10 is a schematic depiction of the physical map and cloning sites (Figure 10A), and the nucleotide sequence (Figure 10B), of the Entry Vector pENTR1A.

-20-

Figure 11 is a schematic depiction of the cloning sites (Figure 11A) and the nucleotide sequence (Figure 11B) of the Entry Vector pENTR2B.

Figure 12 is a schematic depiction of the cloning sites (Figure 12A) and the nucleotide sequence (Figure 12B) of the Entry Vector pENTR3C.

5 Figure 13 is a schematic depiction of the cloning sites (Figure 13A) and the nucleotide sequence (Figure 13B) of the Entry Vector pENTR4.

Figure 14 is a schematic depiction of the cloning sites (Figure 14A) and the nucleotide sequence (Figure 14B) of the Entry Vector pENTR5.

10 Figure 15 is a schematic depiction of the cloning sites (Figure 15A) and the nucleotide sequence (Figure 15B) of the Entry Vector pENTR6.

Figure 16 is a schematic depiction of the cloning sites (Figure 16A) and the nucleotide sequence (Figure 16B) of the Entry Vector pENTR7.

15 Figure 17 is a schematic depiction of the cloning sites (Figure 17A) and the nucleotide sequence (Figure 17B) of the Entry Vector pENTR8.

Figure 18 is a schematic depiction of the cloning sites (Figure 18A) and the nucleotide sequence (Figure 18B) of the Entry Vector pENTR9.

20 Figure 19 is a schematic depiction of the cloning sites (Figure 19A) and the nucleotide sequence (Figure 19B) of the Entry Vector pENTR10.

Figure 20 is a schematic depiction of the cloning sites (Figure 20A) and the nucleotide sequence (Figure 20B) of the Entry Vector pENTR11.

25 Figure 21 is a schematic depiction of the physical map and the Trc expression cassette (Figure 21A) showing the promoter sequences at -35 and at -10 from the initiation codon, and the nucleotide sequence (Figure 21B-D), of Destination Vector pDEST1. This vector may also be referred to as pTrc-DEST1.

Figure 22 is a schematic depiction of the physical map and the His6 expression cassette (Figure 22A) showing the promoter sequences at -35 and at -10 from the initiation codon, and the nucleotide sequence (Figure 22B-D), of Destination Vector pDEST2. This vector may also be referred to as pHis6-DEST2.

5 **Figure 23** is a schematic depiction of the physical map and the GST expression cassette (Figure 23A) showing the promoter sequences at -35 and at -10 from the initiation codon, and the nucleotide sequence (Figure 23B-D), of Destination Vector pDEST3. This vector may also be referred to as pGST-DEST3.

10 **Figure 24** is a schematic depiction of the physical map and the His6-Trx expression cassette (Figure 24A) showing the promoter sequences at -35 and at -10 from the initiation codon and a TEV protease cleavage site, and the nucleotide sequence (Figure 24B-D), of Destination Vector pDEST4. This vector may also be referred to as pTrx-DEST4.

15 **Figure 25** is a schematic depiction of the attR1 and attR2 sites (Figure 25A), the physical map (Figure 25B), and the nucleotide sequence (Figure 25C-D), of Destination Vector pDEST5. This vector may also be referred to as pSPORT(+-)-DEST5.

20 **Figure 26** is a schematic depiction of the attR1 and attR2 sites (Figure 26A), the physical map (Figure 26B), and the nucleotide sequence (Figure 26C-D), of Destination Vector pDEST6. This vector may also be referred to as pSPORT(-)-DEST6.

25 **Figure 27** is a schematic depiction of the attR1 site, CMV promoter, and the physical map (Figure 27A), and the nucleotide sequence (Figure 27B-C), of Destination Vector pDEST7. This vector may also be referred to as pCMV-DEST7.

30 **Figure 28** is a schematic depiction of the attR1 site, baculovirus polyhedrin promoter, and the physical map (Figure 28A), and the nucleotide sequence (Figure 28B-D), of Destination Vector pDEST8. This vector may also be referred to as pFastBac-DEST8.

35 **Figure 29** is a schematic depiction of the attR1 site, Semliki Forest Virus promoter, and the physical map (Figure 29A), and the nucleotide sequence (Figure 29B-E), of Destination Vector pDEST9. This vector may also be referred to as pSFV-DEST9.

5 **Figure 30** is a schematic depiction of the attR1 site, baculovirus polyhedrin promoter, His6 fusion domain, and the physical map (Figure 30A), and the nucleotide sequence (Figure 30B-D), of Destination Vector pDEST10. This vector may also be referred to as pFastBacHT-DEST10.

10 **Figure 31** is a schematic depiction of the attR1 cassette containing a tetracycline-regulated CMV promoter and the physical map (Figure 31A), and the nucleotide sequence (Figure 31B-D), of Destination Vector pDEST11. This vector may also be referred to as pTet-DEST11.

15 **Figure 32** is a schematic depiction of the attR1 site, the start of the mRNA of the CMV promoter, and the physical map (Figure 32A), and the nucleotide sequence (Figure 32B-D), of Destination Vector pDEST12.2. This vector may also be referred to as pCMVneo-DEST12, as pCMV-DEST12, or as pDEST12.

20 **Figure 33** is a schematic depiction of the attR1 site, the λP_L promoter, and the physical map (Figure 33A), and the nucleotide sequence (Figure 33B-C), of Destination Vector pDEST13. This vector may also be referred to as p λP_L -DEST13.

25 **Figure 34** is a schematic depiction of the attR1 site, the T7 promoter, and the physical map (Figure 34A), and the nucleotide sequence (Figure 34B-D), of Destination Vector pDEST14. This vector may also be referred to as pPT7-DEST14.

30 **Figure 35** is a schematic depiction of the attR1 site, the T7 promoter, and the N-terminal GST fusion sequence, and the physical map (Figure 35A), and the nucleotide sequence (Figure 35B-D), of Destination Vector pDEST15. This vector may also be referred to as pT7 GST-DEST15.

35 **Figure 36** is a schematic depiction of the attR1 site, the T7 promoter, and the N-terminal thioredoxin fusion sequence, and the physical map (Figure 36A), and the nucleotide sequence (Figure 36B-D), of Destination Vector pDEST16. This vector may also be referred to as pT7 Trx-DEST16.

40 **Figure 37** is a schematic depiction of the attR1 site, the T7 promoter, and the N-terminal His6 fusion sequence, and the physical map (Figure 37A), and the

nucleotide sequence (Figure 37B-D), of Destination Vector pDEST17. This vector may also be referred to as pT7 His-DEST17.

5 **Figure 38** is a schematic depiction of the attR1 site and the p10 baculovirus promoter, and the physical map (Figure 38A), and the nucleotide sequence (Figure 38B-D), of Destination Vector pDEST18. This vector may also be referred to as pFBp10-DEST18.

10 **Figure 39** is a schematic depiction of the attR1 site, and the 39k baculovirus promoter, and the physical map (Figure 39A), and the nucleotide sequence (Figure 39B-D), of Destination Vector pDEST19. This vector may also be referred to as pFB39k-DEST19.

15 **Figure 40** is a schematic depiction of the attR1 site, the *polh* baculovirus promoter, and the N-terminal GST fusion sequence, and the physical map (Figure 40A), and the nucleotide sequence (Figure 40B-D), of Destination Vector pDEST20. This vector may also be referred to as pFB GST-DEST20.

20 **Figure 41** is a schematic depiction of a 2-hybrid vector with a DNA-binding domain, the attR1 site, and the ADH promoter, and the physical map (Figure 41A), and the nucleotide sequence (Figure 41B-E), of Destination Vector pDEST21. This vector may also be referred to as pDB Leu-DEST21.

25 **Figure 42** is a schematic depiction of a 2-hybrid vector with an activation domain, the attR1 site, and the ADH promoter, and the physical map (Figure 42A), and the nucleotide sequence (Figure 42B-D), of Destination Vector pDEST22. This vector may also be referred to as pPC86-DEST22.

30 **Figure 43** is a schematic depiction of the attR1 and attR2 sites, the T7 promoter, and the C-terminal His6 fusion sequence, and the physical map (Figure 43A), and the nucleotide sequence (Figure 43B-D), of Destination Vector pDEST23. This vector may also be referred to as pC-term-His6-DEST23.

Figure 44 is a schematic depiction of the attR1 and attR2 sites, the T7 promoter, and the C-terminal GST fusion sequence, and the physical map (Figure 44A), and the nucleotide sequence (Figure 44B-D), of Destination Vector pDEST24. This vector may also be referred to as pC-term-GST-DEST24.

Figure 45 is a schematic depiction of the attR1 and attR2 sites, the T7 promoter, and the C-terminal thioredoxin fusion sequence, and the physical map (Figure 45A), and the nucleotide sequence (Figure 45B-D), of Destination Vector pDEST25. This vector may also be referred to as pC-term-Trx-DEST25.

5 Figure 46 is a schematic depiction of the attR1 site, the CMV promoter, and an N-terminal His6 fusion sequence, and the physical map (Figure 46A), and the nucleotide sequence (Figure 46B-D), of Destination Vector pDEST26. This vector may also be referred to as pCMV-SPneo-His-DEST26.

10 Figure 47 is a schematic depiction of the attR1 site, the CMV promoter, and an N-terminal GST fusion sequence, and the physical map (Figure 47A), and the nucleotide sequence (Figure 47B-D), of Destination Vector pDEST27. This vector may also be referred to as pCMV-Spneo-GST-DEST27.

15 Figure 48 is a depiction of the physical map (Figure 48A), the cloning sites (Figure 48B), and the nucleotide sequence (Figure 48C-D), for the attB cloning vector plasmid pEXP501. This vector may also be referred to equivalently herein as pCMV•SPORT6, pCMVSPORT6, and pCMVSport6.

20 Figure 49 is a depiction of the physical map (Figure 49A), and the nucleotide sequence (Figure 49B-C), for the Donor plasmid pDONR201 which donates a kanamycin-resistant vector in the BP Reaction. This vector may also be referred to as pAttPkanr Donor Plasmid, or as pAttPkan Donor Plasmid

Figure 50 is a depiction of the physical map (Figure 50A), and the nucleotide sequence (Figure 50B-C), for the Donor plasmid pDONR202 which donates a kanamycin-resistant vector in the BP Reaction.

25 Figure 51 is a depiction of the physical map (Figure 51A), and the nucleotide sequence (Figure 51B-C), for the Donor plasmid pDONR203 which donates a kanamycin-resistant vector in the BP Reaction.

Figure 52 is a depiction of the physical map (Figure 52A), and the nucleotide sequence (Figure 52B-C), for the Donor plasmid pDONR204 which donates a kanamycin-resistant vector in the BP Reaction.

Figure 53 is a depiction of the physical map (Figure 53A), and the nucleotide sequence (Figure 53B-C), for the Donor plasmid pDONR205 which donates a tetracycline-resistant vector in the BP Reaction.

5 Figure 54 is a depiction of the physical map (Figure 54A), and the nucleotide sequence (Figure 54B-C), for the Donor plasmid pDONR206 which donates a gentamycin-resistant vector in the BP Reaction. This vector may also be referred to as pENTR22 attP Donor Plasmid, pAttPGenr Donor Plasmid, or pAttPgent Donor Plasmid.

10 Figure 55 depicts the attB1 site, and the physical map, of an Entry Clone (pENTR7) of CAT subcloned into the Destination Vector pDEST2 (Figure 22).

Figure 56 depicts the DNA components of Reaction B of the one-tube BxP reaction described in Example 16, pEZC7102 and attB-tet-PCR.

15 Figure 57 is a physical map of the desired product of Reaction B of the one-tube BxP reaction described in Example 16, tetx7102.

Figure 58 is a physical map of the Destination Vector pEZC8402.

20 Figure 59 is a physical map of the expected tet^r subclone product, tetx8402, resulting from the LxR Reaction with tetx7102 (Figure 57) plus pEZC8402 (Figure 58).

Figure 60 is a schematic depiction of the bacteriophage lambda recombination pathways in *E. coli*.

25 Figure 61 is a schematic depiction of the DNA molecules participating in the LR Reaction. Two different co-integrates form during the LR Reaction (only one of which is shown here), depending on whether attL1 and attR1 or attL2 and attR2 are first to recombine. In one aspect, the invention provides directional cloning of a nucleic acid molecule of interest, since the recombination sites react with specificity (attL1 reacts with attR1; attL2 with attR2; attB1 with attP1; and attB2 with attP2). Thus, positioning of the sites allows construction of desired vectors having recombined fragments in the desired orientation.

30 Figure 62 is a depiction of native and fusion protein expression using the recombinational cloning methods and compositions of the invention. In the upper figure depicting native protein expression, all of the translational start signals are

-26-

included between the attB1 and attB2 sites; therefore, these signals must be present in the starting Entry Clone. The lower figure depicts fusion protein expression (here showing expression with both N-terminal and C-terminal fusion tags so that ribosomes read through attB1 and attB2 to create the fusion protein).
5 Unlike native protein expression vectors, N-terminal fusion vectors have their translational start signals upstream of the attB1 site.

10

Figure 63 is a schematic depiction of three GATEWAY™ Cloning System cassettes. Three blunt-ended cassettes are depicted which convert standard expression vectors to Destination Vectors. Each of the depicted cassettes provides amino-terminal fusions in one of three possible reading frames, and each has a distinctive restriction cleavage site as shown.

Figure 64 shows the physical maps of plasmids containing three attR reading frame cassettes, pEYC15101 (reading frame A; Figure 64A), pEYC15102 (reading frame B; Figure 64B), and pEYC15103 (reading frame C; Figure 64C).

15

Figure 65 depicts the attB primers used for amplifying the tet' and amp' genes from pBR322 by the cloning methods of the invention.

Figure 66 is a table listing the results of recombinational cloning of the tet' and amp' PCR products made using the primers shown in Figure 65.

20

Figure 67 is a graph showing the effect of the number of guanines (G's) contained on the 5' end of the PCR primers on the cloning efficiency of PCR products. It is noted, however, that other nucleotides besides guanine (including A, T, C, U or combinations thereof) may be used as 5' extensions on the PCR primers to enhance cloning efficiency of PCR products.

25

Figure 68 is a graph showing a titration of various amounts of attP and attB reactants in the BxP reaction, and the effects on cloning efficiency of PCR products.

30

Figure 69 is a series of graphs showing the effects of various weights (Figure 69A) or moles (Figure 69B) of a 256 bp PCR product on formation of colonies, and on efficiency of cloning of the 256 bp PCR product into a Donor Vector (Figure 69C).

-27-

Figure 70 is a series of graphs showing the effects of various weights (Figure 70A) or moles (Figure 70B) of a 1 kb PCR product on formation of colonies, and on efficiency of cloning of the 1 kb PCR product into a Donor Vector (Figure 70C).

5 Figure 71 is a series of graphs showing the effects of various weights (Figure 71A) or moles (Figure 71B) of a 1.4 kb PCR product on formation of colonies, and on efficiency of cloning of the 1.4 kb PCR product into a Donor Vector (Figure 71C).

10 Figure 72 is a series of graphs showing the effects of various weights (Figure 72A) or moles (Figure 72B) of a 3.4 kb PCR product on formation of colonies, and on efficiency of cloning of the 3.4 kb PCR product into a Donor Vector (Figure 72C).

15 Figure 73 is a series of graphs showing the effects of various weights (Figure 73A) or moles (Figure 73B) of a 4.6 kb PCR product on formation of colonies, and on efficiency of cloning of the 4.6 kb PCR product into a Donor Vector (Figure 73C).

Figure 74 is photograph of an ethidium bromide-stained gel of a titration of a 6.9 kb PCR product in a BxP reaction.

20 Figure 75 is a graph showing the effects of various amounts of a 10.1 kb PCR product on formation of colonies upon cloning of the 10.1 kb PCR product into a Donor Vector.

Figure 76 is photograph of an ethidium bromide-stained gel of a titration of a 10.1 kb PCR product in a BxP reaction.

25 Figure 77 is a table summarizing the results of the PCR product cloning efficiency experiments depicted in Figures 69-74, for PCR fragments ranging in size from 0.256 kb to 6.9 kb.

30 Figure 78 is a depiction of the sequences at the ends of attR Cassettes. Sequences contributed by the Cm^r-ccdB cassette are shown, including the outer ends of the flanking attR sites (boxed). The staggered cleavage sites for Int are indicated in the boxed regions. Following recombination with an Entry Clone, only the outer sequences in attR sites contribute to the resulting attB sites in the

Expression Clone. The underlined sequences at both ends dictate the different reading frames (reading frames A, B, or C, with two alternative reading frame C cassettes depicted) for fusion proteins.

5 **Figure 79** is a depiction of several different attR cassettes (in reading frames A, B, or C) which may provide fusion codons at the amino-terminus of the encoded protein.

10 **Figure 80** illustrates the single-cutting restriction sites in an attR reading frame A cassette of the invention.

15 **Figure 81** illustrates the single-cutting restriction sites in an attR reading frame B cassette of the invention.

20 **Figure 82** illustrates the single-cutting restriction sites in two alternative attR reading frame C cassettes of the invention (Figures 82A and 82B) depicted in Figure 78.

25 **Figure 83** shows the physical map (Figure 83A), and the nucleotide sequence (Figure 83B-C), for an attR reading frame C parent plasmid prfC Parent III, which contains an attR reading frame C cassette of the invention (alternative A in Figures 78 and 82).

30 **Figure 84** is a physical map of plasmid pEZC1301.

20 **Figure 85** is a physical map of plasmid pEZC1313.

25 **Figure 86** is a physical map of plasmid pEZ14032.

30 **Figure 87** is a physical map of plasmid pMAB58.

20 **Figure 88** is a physical map of plasmid pMAB62.

25 **Figure 89** is a depiction of a synthesis reaction using two pairs of homologous primers of the invention.

30 **Figure 90** is a schematic depiction of the physical map (Figure 90A), and the nucleotide sequence (Figure 90B-D), of Destination Vector pDEST28.

20 **Figure 91** is a schematic depiction of the physical map (Figure 91A), and the nucleotide sequence (Figure 91B-D), of Destination Vector pDEST29.

25 **Figure 92** is a schematic depiction of the physical map (Figure 92A), and the nucleotide sequence (Figure 92B-D), of Destination Vector pDEST30.

Figure 93 is a schematic depiction of the physical map (Figure 93A), and the nucleotide sequence (Figure 93B-D), of Destination Vector pDEST31.

Figure 94 is a schematic depiction of the physical map (Figure 94A), and the nucleotide sequence (Figure 94B-E), of Destination Vector pDEST32.

5 Figure 95 is a schematic depiction of the physical map (Figure 95A), and the nucleotide sequence (Figure 95B-D), of Destination Vector pDEST33.

Figure 96 is a schematic depiction of the physical map (Figure 96A), and the nucleotide sequence (Figure 96B-D), of Destination Vector pDEST34.

10 Figure 97 is a depiction of the physical map (Figure 97A), and the nucleotide sequence (Figure 97B-C), for the Donor plasmid pDONR207 which donates a gentamycin-resistant vector in the BP Reaction.

Figure 98 is a schematic depiction of the physical map (Figure 98A), and the nucleotide sequence (Figure 98B-D), of the 2-hybrid vector pMAB85.

15 Figure 99 is a schematic depiction of the physical map (Figure 99A), and the nucleotide sequence (Figure 99B-D), of the 2-hybrid vector pMAB86.

DETAILED DESCRIPTION OF THE INVENTION

20 *Definitions*

In the description that follows, a number of terms used in recombinant DNA technology are utilized extensively. In order to provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided.

25 **Byproduct:** is a daughter molecule (a new clone produced after the second recombination event during the recombinational cloning process) lacking the segment which is desired to be cloned or subcloned.

30 **Cointegrate:** is at least one recombination intermediate nucleic acid molecule of the present invention that contains both parental (starting) molecules. It will usually be linear. In some embodiments it can be circular. RNA and polypeptides may be expressed from cointegrates using an appropriate host cell strain, for example *E. coli* DB3.1 (particularly *E. coli* LIBRARY EFFICIENCY®

DB3.1™ Competent Cells), and selecting for both selection markers found on the cointegrate molecule.

5 **Host:** is any prokaryotic or eukaryotic organism that can be a recipient of the recombinational cloning Product, vector, or nucleic acid molecule of the invention. A "host," as the term is used herein, includes prokaryotic or eukaryotic organisms that can be genetically engineered. For examples of such hosts, see Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1982).

10 **Insert or Inserts:** include the desired nucleic acid segment or a population of nucleic acid segments (segment *A* of Figure 1) which may be manipulated by the methods of the present invention. Thus, the terms Insert(s) are meant to include a particular nucleic acid (preferably DNA) segment or a population of segments. Such Insert(s) can comprise one or more nucleic acid molecules.

15 **Insert Donor:** is one of the two parental nucleic acid molecules (e.g. RNA or DNA) of the present invention which carries the Insert. The Insert Donor molecule comprises the Insert flanked on both sides with recombination sites. The Insert Donor can be linear or circular. In one embodiment of the invention, the Insert Donor is a circular DNA molecule and further comprises a cloning vector sequence outside of the recombination signals (see Figure 1). When a population of Inserts or population of nucleic acid segments are used to make the Insert Donor, a population of Insert Donors results and may be used in accordance with the invention. Examples of such Insert Donor molecules are GATEWAY™ Entry Vectors, which include but are not limited to those Entry Vectors depicted in Figures 10-20, as well as other vectors comprising a gene of interest flanked by one or more *attL* sites (e.g., *attL1*, *attL2*, etc.), or by one or more *attB* sites (e.g., *attB1*, *attB2*, etc.) for the production of library clones.

20 **Product:** is one of the desired daughter molecules comprising the *A* and *D* sequences which is produced after the second recombination event during the recombinational cloning process (see Figure 1). The Product contains the nucleic acid which was to be cloned or subcloned. In accordance with the invention, when a population of Insert Donors are used, the resulting population of Product

molecules will contain all or a portion of the population of Inserts of the Insert Donors and preferably will contain a representative population of the original molecules of the Insert Donors.

Promoter: is a DNA sequence generally described as the 5'-region of a gene, located proximal to the start codon. The transcription of an adjacent DNA segment is initiated at the promoter region. A repressible promoter's rate of transcription decreases in response to a repressing agent. An inducible promoter's rate of transcription increases in response to an inducing agent. A constitutive promoter's rate of transcription is not specifically regulated, though it can vary under the influence of general metabolic conditions.

Recognition sequence: Recognition sequences are particular sequences which a protein, chemical compound, DNA, or RNA molecule (e.g., restriction endonuclease, a modification methylase, or a recombinase) recognizes and binds. In the present invention, a recognition sequence will usually refer to a recombination site. For example, the recognition sequence for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See Figure 1 of Sauer, B., *Current Opinion in Biotechnology* 5:521-527 (1994). Other examples of recognition sequences are the *attB*, *attP*, *attL*, and *attR* sequences which are recognized by the recombinase enzyme λ Integrase. *attB* is an approximately 25 base pair sequence containing two 9 base pair core-type Int binding sites and a 7 base pair overlap region. *attP* is an approximately 240 base pair sequence containing core-type Int binding sites and arm-type Int binding sites as well as sites for auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). See Landy, *Current Opinion in Biotechnology* 3:699-707 (1993). Such sites may also be engineered according to the present invention to enhance production of products in the methods of the invention. When such engineered sites lack the P1 or H1 domains to make the recombination reactions irreversible (e.g., *attR'* or *attP'*), such sites may be designated *attR'* or *attP'* to show that the domains of these sites have been modified in some way.

5 **Recombination proteins:** include excisive or integrative proteins, enzymes, co-factors or associated proteins that are involved in recombination reactions involving one or more recombination sites, which may be wild-type proteins (*See Landy, Current Opinion in Biotechnology 3:699-707 (1993)*), or mutants, derivatives (e.g., fusion proteins containing the recombination protein sequences or fragments thereof), fragments, and variants thereof.

10 **Recombination site:** is a recognition sequence on a DNA molecule participating in an integration/recombination reaction by the recombinational cloning methods of the invention. Recombination sites are discrete sections or segments of DNA on the participating nucleic acid molecules that are recognized and bound by a site-specific recombination protein during the initial stages of integration or recombination. For example, the recombination site for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. *See Figure 1 of Sauer, B., Curr. Opin. Biotech. 5:521-527 (1994).* Other examples of recognition sequences include the *attB*, *attP*, *attL*, and *attR* sequences described herein, and mutants, fragments, variants and derivatives thereof, which are recognized by the recombination protein λ Int and by the auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). *See Landy, Curr. Opin. Biotech. 3:699-707 (1993).*

15 **Recombinational Cloning:** is a method described herein, whereby segments of nucleic acid molecules or populations of such molecules are exchanged, inserted, replaced, substituted or modified, *in vitro* or *in vivo*. By "in vitro" and "in vivo" herein is meant recombinational cloning that is carried out outside of host cells (e.g., in cell-free systems) or inside of host cells (e.g., using recombinant proteins expressed by host cells), respectively.

20 **Repression cassette:** is a nucleic acid segment that contains a repressor or a Selectable marker present in the subcloning vector.

25 **Selectable marker:** is a DNA segment that allows one to select for or against a molecule (e.g., a replicon) or a cell that contains it, often under particular conditions. These markers can encode an activity, such as, but not limited to,

production of RNA, peptide, or protein, or can provide a binding site for RNA, peptides, proteins, inorganic and organic compounds or compositions and the like. Examples of Selectable markers include but are not limited to: (1) DNA segments that encode products which provide resistance against otherwise toxic compounds (e.g., antibiotics); (2) DNA segments that encode products which are otherwise lacking in the recipient cell (e.g., tRNA genes, auxotrophic markers); (3) DNA segments that encode products which suppress the activity of a gene product; (4) DNA segments that encode products which can be readily identified (e.g., phenotypic markers such as β -galactosidase, green fluorescent protein (GFP), and cell surface proteins); (5) DNA segments that bind products which are otherwise detrimental to cell survival and/or function; (6) DNA segments that otherwise inhibit the activity of any of the DNA segments described in Nos. 1-5 above (e.g., antisense oligonucleotides); (7) DNA segments that bind products that modify a substrate (e.g. restriction endonucleases); (8) DNA segments that can be used to isolate or identify a desired molecule (e.g. specific protein binding sites); (9) DNA segments that encode a specific nucleotide sequence which can be otherwise non-functional (e.g., for PCR amplification of subpopulations of molecules); (10) DNA segments, which when absent, directly or indirectly confer resistance or sensitivity to particular compounds; (11) DNA segments that encode products which are toxic in recipient cells; (12) DNA segments that inhibit replication, partition or heritability of nucleic acid molecules that contain them; and/or (13) DNA segments that encode conditional replication functions, e.g., replication in certain hosts or host cell strains or under certain environmental conditions (e.g., temperature, nutritional conditions, etc.).

Selection scheme: is any method which allows selection, enrichment, or identification of a desired Product or Product(s) from a mixture containing an Entry Clone or Vector, a Destination Vector, a Donor Vector, an Expression Clone or Vector, any intermediates (e.g. a Cointegrate or a replicon), and/or Byproducts. The selection schemes of one preferred embodiment have at least two components that are either linked or unlinked during recombinational cloning. One component is a Selectable marker. The other component controls the expression *in vitro* or *in vivo* of the Selectable marker, or survival of the cell (or

5

10

15

20

25

30

the nucleic acid molecule, e.g., a replicon) harboring the plasmid carrying the Selectable marker. Generally, this controlling element will be a repressor or inducer of the Selectable marker, but other means for controlling expression or activity of the Selectable marker can be used. Whether a repressor or activator is used will depend on whether the marker is for a positive or negative selection, and the exact arrangement of the various DNA segments, as will be readily apparent to those skilled in the art. A preferred requirement is that the selection scheme results in selection of or enrichment for only one or more desired Products. As defined herein, selecting for a DNA molecule includes (a) selecting or enriching for the presence of the desired DNA molecule, and (b) selecting or enriching against the presence of DNA molecules that are not the desired DNA molecule.

In one embodiment, the selection schemes (which can be carried out in reverse) will take one of three forms, which will be discussed in terms of Figure 1. The first, exemplified herein with a Selectable marker and a repressor therefore, selects for molecules having segment *D* and lacking segment *C*. The second selects against molecules having segment *C* and for molecules having segment *D*. Possible embodiments of the second form would have a DNA segment carrying a gene toxic to cells into which the *in vitro* reaction products are to be introduced. A toxic gene can be a DNA that is expressed as a toxic gene product (a toxic protein or RNA), or can be toxic in and of itself. (In the latter case, the toxic gene is understood to carry its classical definition of "heritable trait".)

Examples of such toxic gene products are well known in the art, and include, but are not limited to, restriction endonucleases (e.g., *Dpn*I), apoptosis-related genes (e.g. ASK1 or members of the bcl-2/ced-9 family), retroviral genes including those of the human immunodeficiency virus (HIV), defensins such as NP-1, inverted repeats or paired palindromic DNA sequences, bacteriophage lytic genes such as those from Φ X174 or bacteriophage T4; antibiotic sensitivity genes such as *rpsL*, antimicrobial sensitivity genes such as *pheS*, plasmid killer genes, eukaryotic transcriptional vector genes that produce a gene product toxic to bacteria, such as GATA-1, and genes that kill hosts in the absence of a suppressing function, e.g., *kicB*, *ccdB*, Φ X174 *E* (Liu, Q. et al., *Curr. Biol.*

8:1300-1309 (1998)), and other genes that negatively affect replicon stability and/or replication. A toxic gene can alternatively be selectable *in vitro*, e.g., a restriction site.

Many genes coding for restriction endonucleases operably linked to inducible promoters are known, and may be used in the present invention. See, e.g. U.S. Patent Nos. 4,960,707 (*DpnI* and *DpnII*); 5,000,333, 5,082,784 and 5,192,675 (*KpnI*); 5,147,800 (*NgoAIII* and *NgoAI*); 5,179,015 (*FspI* and *HaeIII*); 5,200,333 (*HaeII* and *TaqI*); 5,248,605 (*HpaII*); 5,312,746 (*Clal*); 5,231,021 and 5,304,480 (*XbaI* and *XbaII*); 5,334,526 (*AluI*); 5,470,740 (*NsiI*); 5,534,428 (*SstI/SacI*); 5,202,248 (*NcoI*); 5,139,942 (*NdeI*); and 5,098,839 (*PacI*). See also Wilson, G.G., *Nucl. Acids Res.* 19:2539-2566 (1991); and Lunnen, K.D., *et al.*, *Gene* 74:25-32 (1988).

In the second form, segment *D* carries a Selectable marker. The toxic gene would eliminate transformants harboring the Vector Donor, Cointegrate, and Byproduct molecules, while the Selectable marker can be used to select for cells containing the Product and against cells harboring only the Insert Donor.

The third form selects for cells that have both segments *A* and *D* in *cis* on the same molecule, but not for cells that have both segments in *trans* on different molecules. This could be embodied by a Selectable marker that is split into two inactive fragments, one each on segments *A* and *D*.

The fragments are so arranged relative to the recombination sites that when the segments are brought together by the recombination event, they reconstitute a functional Selectable marker. For example, the recombinational event can link a promoter with a structural nucleic acid molecule (e.g., a gene), can link two fragments of a structural nucleic acid molecule, or can link nucleic acid molecules that encode a heterodimeric gene product needed for survival, or can link portions of a replicon.

Site-specific recombinase: is a type of recombinase which typically has at least the following four activities (or combinations thereof): (1) recognition of one or two specific nucleic acid sequences; (2) cleavage of said sequence or sequences; (3) topoisomerase activity involved in strand exchange; and (4) ligase

activity to reseal the cleaved strands of nucleic acid. See Sauer, B., *Current Opinions in Biotechnology* 5:521-527 (1994). Conservative site-specific recombination is distinguished from homologous recombination and transposition by a high degree of sequence specificity for both partners. The strand exchange mechanism involves the cleavage and rejoining of specific DNA sequences in the absence of DNA synthesis (Landy, A. (1989) *Ann. Rev. Biochem.* 58:913-949).

Subcloning vector: is a cloning vector comprising a circular or linear nucleic acid molecule which includes preferably an appropriate replicon. In the present invention, the subcloning vector (segment *D* in Figure 1) can also contain functional and/or regulatory elements that are desired to be incorporated into the final product to act upon or with the cloned DNA Insert (segment *A* in Figure 1). The subcloning vector can also contain a Selectable marker (preferably DNA).

Vector: is a nucleic acid molecule (preferably DNA) that provides a useful biological or biochemical property to an Insert. Examples include plasmids, phages, autonomously replicating sequences (ARS), centromeres, and other sequences which are able to replicate or be replicated *in vitro* or in a host cell, or to convey a desired nucleic acid segment to a desired location within a host cell. A Vector can have one or more restriction endonuclease recognition sites at which the sequences can be cut in a determinable fashion without loss of an essential biological function of the vector, and into which a nucleic acid fragment can be spliced in order to bring about its replication and cloning. Vectors can further provide primer sites, e.g., for PCR, transcriptional and/or translational initiation and/or regulation sites, recombinational signals, replicons, Selectable markers, etc. Clearly, methods of inserting a desired nucleic acid fragment which do not require the use of homologous recombination, transpositions or restriction enzymes (such as, but not limited to, UDG cloning of PCR fragments (U.S. Patent No. 5,334,575, entirely incorporated herein by reference), T:A cloning, and the like) can also be applied to clone a fragment into a cloning vector to be used according to the present invention. The cloning vector can further contain one or more selectable markers suitable for use in the identification of cells transformed with the cloning vector.

5 **Vector Donor:** is one of the two parental nucleic acid molecules (e.g., RNA or DNA) of the present invention which carries the DNA segments comprising the DNA vector which is to become part of the desired Product. The Vector Donor comprises a subcloning vector *D* (or it can be called the cloning vector if the Insert Donor does not already contain a cloning vector (e.g., for PCR fragments containing *attB* sites; see below)) and a segment *C* flanked by recombination sites (see Figure 1). Segments *C* and/or *D* can contain elements that contribute to selection for the desired Product daughter molecule, as described above for selection schemes. The recombination signals can be the same or different, and can be acted upon by the same or different recombinases. In addition, the Vector Donor can be linear or circular. Examples of such Vector Donor molecules include GATEWAY™ Destination Vectors, which include but are not limited to those Destination Vectors depicted in Figures 21-47 and 90-96.

10

15 **Primer:** refers to a single stranded or double stranded oligonucleotide that is extended by covalent bonding of nucleotide monomers during amplification or polymerization of a nucleic acid molecule (e.g. a DNA molecule). In a preferred aspect, a primer comprises one or more recombination sites or portions of such recombination sites. Portions of recombination sites comprise at least 2 bases (or basepairs, abbreviated herein as "bp"), at least 5-200 bases, at least 10-100 bases, at least 15-75 bases, at least 15-50 bases, at least 15-25 bases, or at least 16-25 bases, of the recombination sites of interest, as described in further detail below and in the Examples. When using portions of recombination sites, the missing portion of the recombination site may be provided as a template by the newly synthesized nucleic acid molecule. Such recombination sites may be located within and/or at one or both termini of the primer. Preferably, additional sequences are added to the primer adjacent to the recombination site(s) to enhance or improve recombination and/or to stabilize the recombination site during recombination. Such stabilization sequences may be any sequences (preferably G/C rich sequences) of any length. Preferably, such sequences range in size from 1 to about 1000 bases, 1 to about 500 bases, and 1 to about 100 bases, 1 to about 60 bases, 1 to about 25, 1 to about 10, 2 to about 10 and preferably about 4 bases.

20

25

30

Preferably, such sequences are greater than 1 base in length and preferably greater than 2 bases in length.

Template: refers to double stranded or single stranded nucleic acid molecules which are to be amplified, synthesized or sequenced. In the case of double stranded molecules, denaturation of its strands to form a first and a second strand is preferably performed before these molecules will be amplified, synthesized or sequenced, or the double stranded molecule may be used directly as a template. For single stranded templates, a primer complementary to a portion of the template is hybridized under appropriate conditions and one or more polypeptides having polymerase activity (e.g. DNA polymerases and/or reverse transcriptases) may then synthesize a nucleic acid molecule complementary to all or a portion of said template. Alternatively, for double stranded templates, one or more promoters may be used in combination with one or more polymerases to make nucleic acid molecules complementary to all or a portion of the template. The newly synthesized molecules, according to the invention, may be equal or shorter in length than the original template. Additionally, a population of nucleic acid templates may be used during synthesis or amplification to produce a population of nucleic acid molecules typically representative of the original template population.

Adapter: is an oligonucleotide or nucleic acid fragment or segment (preferably DNA) which comprises one or more recombination sites (or portions of such recombination sites) which in accordance with the invention can be added to a circular or linear Insert Donor molecule as well as other nucleic acid molecules described herein. When using portions of recombination sites, the missing portion may be provided by the Insert Donor molecule. Such adapters may be added at any location within a circular or linear molecule, although the adapters are preferably added at or near one or both termini of a linear molecule. Preferably, adapters are positioned to be located on both sides (flanking) a particular nucleic acid molecule of interest. In accordance with the invention, adapters may be added to nucleic acid molecules of interest by standard recombinant techniques (e.g. restriction digest and ligation). For example, adapters may be added to a circular molecule by first digesting the molecule with

5

10

an appropriate restriction enzyme, adding the adapter at the cleavage site and reforming the circular molecule which contains the adapter(s) at the site of cleavage. In other aspects, adapters may be added by homologous recombination, by integration of RNA molecules, and the like. Alternatively, adapters may be ligated directly to one or more and preferably both termini of a linear molecule thereby resulting in linear molecule(s) having adapters at one or both termini. In one aspect of the invention, adapters may be added to a population of linear molecules, (e.g. a cDNA library or genomic DNA which has been cleaved or digested) to form a population of linear molecules containing adapters at one and preferably both termini of all or substantial portion of said population.

15

20

25

Adapter-Primer: is primer molecule which comprises one or more recombination sites (or portions of such recombination sites) which in accordance with the invention can be added to a circular or linear nucleic acid molecule described herein. When using portions of recombination sites, the missing portion may be provided by a nucleic acid molecule (e.g., an adapter) of the invention. Such adapter-primers may be added at any location within a circular or linear molecule, although the adapter-primers are preferably added at or near one or both termini of a linear molecule. Examples of such adapter-primers and the use thereof in accordance with the methods of the invention are shown in Example 25 herein. Such adapter-primers may be used to add one or more recombination sites or portions thereof to circular or linear nucleic acid molecules in a variety of contexts and by a variety of techniques, including but not limited to amplification (e.g., PCR), ligation (e.g., enzymatic or chemical/synthetic ligation), recombination (e.g., homologous or non-homologous (illegitimate) recombination) and the like.

30

Library: refers to a collection of nucleic acid molecules (circular or linear). In one embodiment, a library may comprise a plurality (*i.e.*, two or more) of DNA molecules, which may or may not be from a common source organism, organ, tissue, or cell. In another embodiment, a library is representative of all or a portion or a significant portion of the DNA content of an organism (a "genomic" library), or a set of nucleic acid molecules representative of all or a portion or a significant portion of the expressed nucleic acid molecules (a cDNA library) in a

cell, tissue, organ or organism. A library may also comprise random sequences made by *de novo* synthesis, mutagenesis of one or more sequences and the like. Such libraries may or may not be contained in one or more vectors.

5 **Amplification:** refers to any *in vitro* method for increasing a number of copies of a nucleotide sequence with the use of a polymerase. Nucleic acid amplification results in the incorporation of nucleotides into a DNA and/or RNA molecule or primer thereby forming a new molecule complementary to a template. The formed nucleic acid molecule and its template can be used as templates to synthesize additional nucleic acid molecules. As used herein, one amplification reaction may consist of many rounds of replication. DNA amplification reactions include, for example, polymerase chain reaction (PCR). One PCR reaction may consist of 5-100 "cycles" of denaturation and synthesis of a DNA molecule.

10 **Oligonucleotide:** refers to a synthetic or natural molecule comprising a covalently linked sequence of nucleotides which are joined by a phosphodiester bond between the 3' position of the deoxyribose or ribose of one nucleotide and the 5' position of the deoxyribose or ribose of the adjacent nucleotide. This term may be used interchangeably herein with the terms "nucleic acid molecule" and "polynucleotide," without any of these terms necessarily indicating any particular length of the nucleic acid molecule to which the term specifically refers.

15 **Nucleotide:** refers to a base-sugar-phosphate combination. Nucleotides are monomeric units of a nucleic acid molecule (DNA and RNA). The term nucleotide includes ribonucleoside triphosphates ATP, UTP, CTG, GTP and deoxyribonucleoside triphosphates such as dATP, dCTP, dITP, dUTP, dGTP, dTTP, or derivatives thereof. Such derivatives include, for example, [α S]dATP, 7-deaza-dGTP and 7-deaza-dATP. The term nucleotide as used herein also refers to dideoxyribonucleoside triphosphates (ddNTPs) and their derivatives. Illustrated examples of dideoxyribonucleoside triphosphates include, but are not limited to, ddATP, ddCTP, ddGTP, ddITP, and ddTTP. According to the present invention, a "nucleotide" may be unlabeled or detectably labeled by well known techniques. Detectable labels include, for example, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels and enzyme labels.

5 **Hybridization:** The terms "hybridization" and "hybridizing" refers to base pairing of two complementary single-stranded nucleic acid molecules (RNA and/or DNA) to give a double stranded molecule. As used herein, two nucleic acid molecules may be hybridized, although the base pairing is not completely complementary. Accordingly, mismatched bases do not prevent hybridization of two nucleic acid molecules provided that appropriate conditions, well known in the art, are used. In some aspects, hybridization is said to be under "stringent conditions." By "stringent conditions" as used herein is meant overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150 mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 g/ml denatured, sheared salmon 10 sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

10 Other terms used in the fields of recombinant DNA technology and molecular and cell biology as used herein will be generally understood by one of ordinary skill in the applicable arts.

15 *Overview*

20 Two reactions constitute the recombinational cloning system of the present invention, referred to herein as the "GATEWAY™ Cloning System," as depicted generally in Figure 1. The first of these reactions, the LR Reaction (Figure 2), which may also be referred to interchangeably herein as the Destination Reaction, is the main pathway of this system. The LR Reaction is a recombination reaction between an Entry vector or clone and a Destination Vector, mediated by a cocktail of recombination proteins such as the GATEWAY™ LR Clonase™ Enzyme Mix described herein. This reaction 25 transfers nucleic acid molecules of interest (which may be genes, cDNAs, cDNA libraries, or fragments thereof) from the Entry Clone to an Expression Vector, to create an Expression Clone.

30 The sites labeled L, R, B, and P are respectively the attL, attR, attB, and attP recombination sites for the bacteriophage λ recombination proteins that constitute the Clonase cocktail (referred to herein variously as "Clonase" or

“GATEWAY™ LR Clonase™ Enzyme Mix” (for recombination protein mixtures mediating attL x attR recombination reactions, as described herein) or “GATEWAY™ BP Clonase™ Enzyme Mix” (for recombination protein mixtures mediating attB x attP recombination reactions, as described herein)). The Recombinational Cloning reactions are equivalent to concerted, highly specific, cutting and ligation reactions. Viewed in this way, the recombination proteins cut to the left and right of the nucleic acid molecule of interest in the Entry Clone and ligate it into the Destination vector, creating a new Expression Clone.

The nucleic acid molecule of interest in an Expression Clone is flanked by the small attB1 and attB2 sites. The orientation and reading frame of the nucleic acid molecule of interest are maintained throughout the subcloning, because attL1 reacts only with attR1, and attL2 reacts only with attR2. Likewise, attB1 reacts only with attP1, and attB2 reacts only with attP2. Thus, the invention also relates to methods of controlled or directional cloning using the recombination sites of the invention (or portions thereof), including variants, fragments, mutants and derivatives thereof which may have altered or enhanced specificity. The invention also relates more generally to any number of recombination site partners or pairs (where each recombination site is specific for and interacts with its corresponding recombination site). Such recombination sites are preferably made by mutating or modifying the recombination site to provide any number of necessary specificities (e.g., attB1-10, attP1-10, attL1-10, attR1-10, etc.), non-limiting examples of which are described in detail in the Examples herein.

When an aliquot from the recombination reaction is transformed into host cells (e.g., *E. coli*) and spread on plates containing an appropriate selection agent, e.g., an antibiotic such as ampicillin with or without methicillin, cells that take up the desired clone form colonies. The unreacted Destination Vector does not give ampicillin-resistant colonies, even though it carries the ampicillin-resistance gene, because it contains a toxic gene, e.g., *ccdB*. Thus selection for ampicillin resistance selects for *E. coli* cells that carry the desired product, which usually comprise >90% of the colonies on the ampicillin plate.

To participate in the Recombinational (or “GATEWAY™”) Cloning Reaction, a nucleic acid molecule of interest first may be cloned into an Entry

5 Vector, creating an Entry Clone. Multiple options are available for creating Entry Clones, including: cloning of PCR sequences with terminal attB recombination sites into Entry Vectors; using the GATEWAY™ Cloning System recombination reaction; transfer of genes from libraries prepared in GATEWAY™ Cloning System vectors by recombination into Entry Vectors; and cloning of restriction enzyme-generated fragments and PCR fragments into Entry Vectors by standard recombinant DNA methods. These approaches are discussed in further detail herein.

10 A key advantage of the GATEWAY™ Cloning System is that a nucleic acid molecule of interest (or even a population of nucleic acid molecules of interest) present as an Entry Clone can be subcloned in parallel into one or more Destination Vectors in a simple reactions for anywhere from about 30 seconds to about 60 minutes (preferably about 1-60 minutes, about 1-45 minutes, about 1-30 minutes, about 2-60 minutes, about 2-45 minutes, about 2-30 minutes, about 1-2 minutes, about 30-60 minutes, about 45-60 minutes, or about 30-45 minutes).
15 Longer reaction times (e.g., 2-24 hours, or overnight) may increase recombination efficiency, particularly where larger nucleic acid molecules are used, as described in the Examples herein. Moreover, a high percentage of the colonies obtained carry the desired Expression Clone. This process is illustrated schematically in
20 Figure 3, which shows an advantage of the invention in which the molecule of interest can be moved simultaneously or separately into multiple Destination Vectors. In the LR Reaction, one or both of the nucleic acid molecules to be recombined may have any topology (e.g., linear, relaxed circular, nicked circular, supercoiled, etc.), although one or both are preferably linear.

25 The second major pathway of the GATEWAY™ Cloning System is the BP Reaction (Figure 4), which may also be referred to interchangeably herein as the Entry Reaction or the Gateward Reaction. The BP Reaction may recombine an Expression Clone with a Donor Plasmid (the counterpart of the byproduct in Figure 2). This reaction transfers the nucleic acid molecule of interest (which may have any of a variety of topologies, including linear, coiled, supercoiled, etc.) in the Expression Clone into an Entry Vector, to produce a new Entry Clone. Once this nucleic acid molecule of interest is cloned into an Entry

-44-

Vector, it can be transferred into new Expression Vectors, through the LR Reaction as described above. In the BP Reaction, one or both of the nucleic acid molecules to be recombined may have any topology (*e.g.*, linear, relaxed circular, nicked circular, supercoiled, etc.), although one or both are preferably linear.

5 A useful variation of the BP Reaction permits rapid cloning and expression of products of amplification (*e.g.*, PCR) or nucleic acid synthesis. Amplification (*e.g.*, PCR) products synthesized with primers containing terminal 25 bp attB sites serve as efficient substrates for the Gateward Cloning reaction. Such amplification products may be recombined with a Donor Vector to produce an Entry Clone (see
10 Figure 7). The result is an Entry Clone containing the amplification fragment. Such Entry Clones can then be recombined with Destination Vectors -- through the LR Reaction -- to yield Expression Clones of the PCR product.

15 Additional details of the LR Reaction are shown in Figure 5A. The GATEWAY™ LR Clonase™ Enzyme Mix that mediates this reaction contains lambda recombination proteins Int (Integrase), Xis (Excisionase), and IHF (Integration Host Factor). In contrast, the GATEWAY™ BP Clonase™ Enzyme Mix, which mediates the BP Reaction (Figure 5B), comprises Int and IHF alone.

20 The recombination (att) sites of each vector comprise two distinct segments, donated by the parental vectors. The staggered lines dividing the two portions of each att site, depicted in Figures 5A and 5B, represent the seven-base staggered cut produced by Int during the recombination reactions. This structure is seen in greater detail in Figure 6, which displays the attB recombination sequences of an Expression Clone, generated by recombination between the attL1 and attL2 sites of an Entry Clone and the attR1 and attR2 sites of a Destination
25 Vector.

30 The nucleic acid molecule of interest in the Expression Clone is flanked by attB sites: attB1 to the left (amino terminus) and attB2 to the right (carboxy terminus). The bases in attB1 to the left of the seven-base staggered cut produced by Int are derived from the Destination vector, and the bases to the right of the staggered cut are derived from the Entry Vector (see Figure 6). Note that the sequence is displayed in triplets corresponding to an open reading frame. If the reading frame of the nucleic acid molecule of interest cloned in the Entry Vector

-45-

is in phase with the reading frame shown for attB1, amino-terminal protein fusions can be made between the nucleic acid molecule of interest and any GATEWAY™ Cloning System Destination Vector encoding an amino-terminal fusion domain. Entry Vectors and Destination Vectors that enable cloning in all three reading frames are described in more detail herein, particularly in the Examples.

5

The LR Reaction allows the transfer of a desired nucleic acid molecule of interest into new Expression Vectors by recombining a Entry Clone with various Destination Vectors. To participate in the LR or Destination Reaction, however, a nucleic acid molecule of interest preferably is first converted to a Entry Clone. Entry Clones can be made in a number of ways, as shown in Figure 7.

10

One approach is to clone the nucleic acid molecule of interest into one or more of the Entry Vectors, using standard recombinant DNA methods, with restriction enzymes and ligase. The starting DNA fragment can be generated by restriction enzyme digestion or as a PCR product. The fragment is cloned between the attL1 and attL2 recombination sites in the Entry Vector. Note that a toxic or "death" gene (e.g., ccdB), provided to minimize background colonies from incompletely digested Entry Vector, must be excised and replaced by the nucleic acid molecule of interest.

15

A second approach to making an Entry Clone (Figure 7) is to make a library (genomic or cDNA) in an Entry Vector, as described in detail herein. Such libraries may then be transferred into Destination Vectors for expression screening, for example in appropriate host cells such as yeast cells or mammalian cells.

20

A third approach to making Entry Clones (Figure 7) is to use Expression Clones obtained from cDNA molecules or libraries prepared in Expression Vectors. Such cDNAs or libraries, flanked by attB sites, can be introduced into a Entry Vector by recombination with a Donor Vector via the BP Reaction. If desired, an entire Expression Clone library can be transferred into the Entry Vector through the BP Reaction. Expression Clone cDNA libraries may also be constructed in a variety of prokaryotic and eukaryotic GATEWAY™-modified vectors (e.g., the pEXP501 Expression Vector (see Figure 48), and 2-hybrid and

25

30

attB library vectors), as described in detail herein, particularly in the Examples below.

A fourth, and potentially most versatile, approach to making an Entry Clone (Figure 7) is to introduce a sequence for a nucleic acid molecule of interest into an Entry Vector by amplification (e.g., PCR) fragment cloning. This method is diagramed in Figure 8. The DNA sequence first is amplified (for example, with PCR) as outlined in detail below and in the Examples herein, using primers containing one or more bp, two or more bp, three or more bp, four or more bp, five or more bp, preferably six or more bp, more preferably 6-25 bp (particularly 5 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25) bp of the attB nucleotide sequences (such as, but not limited to, those depicted in Figure 9), and optionally one or more, two or more, three or more, four or more, and most preferably four 10 or five or more additional terminal nucleotide bases which preferably are guanines. The PCR product then may be converted to a Entry Clone by performing a BP Reaction, in which the attB-PCR product recombines with a Donor Vector 15 containing one or more attP sites. Details of this approach and protocols for PCR fragment subcloning are provided in Examples 8 and 21-25.

A variety of Entry Clones may be produced by these methods, providing a wide array of cloning options; a number of specific Entry Vectors are also 20 available commercially from Life Technologies, Inc. (Rockville, MD). The Examples herein provide a more in-depth description of selected Entry Vectors and details of their cloning sites. Choosing the optimal Entry Vector for a particular application is discussed in Example 4.

Entry Vectors and Destination Vectors should be constructed so that the 25 amino-terminal region of a nucleic acid molecule of interest (e.g., a gene, cDNA library or insert, or fragment thereof) will be positioned next to the attL1 site. Entry Vectors preferably contain the rrnB transcriptional terminator upstream of the attL1 site. This sequence ensures that expression of cloned nucleic acid molecules of interest is reliably "off" in *E. coli*, so that even toxic genes can be successfully cloned. Thus, Entry Clones may be designed to be transcriptionally 30 silent. Note also that Entry Vectors, and hence Entry Clones, may contain the kanamycin antibiotic resistance (kan^r) gene to facilitate selection of host cells

containing Entry Clones after transformation. In certain applications, however, Entry Clones may contain other selection markers, including but not limited to a gentamycin resistance (*gen*^r) or tetracycline resistance (*tet*^r) gene, to facilitate selection of host cells containing Entry Clones after transformation.

Once a nucleic acid molecule of interest has been cloned into an Entry Vector, it may be moved into a Destination Vector. The upper right portion of Figure 5A shows a schematic of a Destination Vector. The thick arrow represents some function (often transcription or translation) that will act on the nucleic acid molecule of interest in the clone. During the recombination reaction, the region between the attR1 and attR2 sites, including a toxic or "death" gene (e.g., *ccdB*), is replaced by the DNA segment from the Entry Clone. Selection for recombinants that have acquired the ampicillin resistance (*amp*^r) gene (carried on the Destination Vector) and that have also lost the death gene ensures that a high percentage (usually >90%) of the resulting colonies will contain the correct insert.

To move a nucleic acid molecule of interest into a Destination Vector, the Destination Vector is mixed with the Entry Clone comprising the desired nucleic acid molecule of interest, a cocktail of recombination proteins (e.g., GATEWAY™ LR Clonase™ Enzyme Mix) is added, the mixture is incubated (preferably at about 25°C for about 60 minutes, or longer under certain circumstances, e.g. for transfer of large nucleic acid molecules, as described below) and any standard host cell (including bacterial cells such as *E. coli*; animal cells such as insect cells, mammalian cells, nematode cells and the like; plant cells; and yeast cells) strain is transformed with the reaction mixture. The host cell used will be determined by the desired selection (e.g., *E. coli* DB3.1, available commercially from Life Technologies, Inc., allows survival of clones containing the *ccdB* death gene, and thus can be used to select for cointegrate molecules -- i.e., molecules that are hybrids between the Entry Clone and Destination Vector). The Examples below provide further details and protocols for use of Entry and Destination Vectors in transferring nucleic acid molecules of interest and expressing RNAs or polypeptides encoded by these nucleic acid molecules in a variety of host cells.

The cloning system of the invention therefore offers multiple advantages:

- Once a nucleic acid molecule of interest is cloned into the GATEWAY™ Cloning System, it can be moved into and out of other vectors with complete fidelity of reading frame and orientation. That is, since the reactions proceed whereby attL1 on the Entry Clone recombines with attR1 on the Destination Vector, the directionality of the nucleic acid molecule of interest is maintained or may be controlled upon transfer from the Entry Clone into the Destination Vector. Hence, the GATEWAY™ Cloning System provides a powerful and easy method of directional cloning of nucleic acid molecule of interest.
- One-step cloning or subcloning: Mix the Entry Clone and the Destination Vector with Clonase, incubate, and transform.
- Clone PCR products readily by *in vitro* recombination, by adding attB sites to PCR primers. Then directly transfer these Entry Clones into Destination Vectors. This process may also be carried out in one step (see Examples below).
- Powerful selections give high reliability: >90% (and often >99%) of the colonies contain the desired DNA in its new vector.
- One-step conversion of existing standard vectors into GATEWAY™ Cloning System vectors.
- Ideal for large vectors or those with few cloning sites.
- Recombination sites are short (25 bp), and may be engineered to contain no stop codons or secondary structures.
- Reactions may be automated, for high-throughput applications (e.g., for diagnostic purposes or for therapeutic candidate screening).
- The reactions are economical: 0.3 µg of each DNA; no restriction enzymes, phosphatase, ligase, or gel purification. Reactions work well with miniprep DNA.
- Transfer multiple clones, and even libraries, into one or more Destination Vectors, in a single experiment.
- A variety of Destination Vectors may be produced, for applications including, but not limited to:

- Protein expression in *E. coli*: native proteins; fusion proteins with GST, His6, thioredoxin, etc., for purification, or one or more epitope tags; any promoter useful in expressing proteins in *E. coli* may be used, such as ptrc, λP_L , and T7 promoters.
- 5 • Protein expression in eukaryotic cells: CMV promoter, baculovirus (with or without His6 tag), Semliki Forest virus, Tet regulation.
- DNA sequencing (all lac primers), RNA probes, phagemids (both strands)
- 10 • A variety of Entry Vectors (for recombinational cloning entry by standard recombinant DNA methods) may be produced:
 - Strong transcription stop just upstream, for genes toxic to *E. coli*.
 - Three reading frames.
 - With or without TEV protease cleavage site.
 - Motifs for prokaryotic and / or eukaryotic translation.
 - 15 • Compatible with commercial cDNA libraries.
- Expression Clone cDNA (attB) libraries, for expression screening, including 2-hybrid libraries and phage display libraries, may also be constructed.

Recombination Site Sequences

20 In one aspect, the invention relates to nucleic acid molecules, which may or may not be isolated nucleic acid molecules, comprising one or more nucleotide sequences encoding one or more recombination sites or portions thereof. In particular, this aspect of the invention relates to such nucleic acid molecules comprising one or more nucleotide sequences encoding attB, attP, attL, or attR, or portions of these recombination site sequences. The invention also relates to mutants, derivatives, and fragments of such nucleic acid molecules. Unless otherwise indicated, all nucleotide sequences that may have been determined by sequencing a DNA molecule herein were determined using manual or automated DNA sequencing, such as dideoxy sequencing, according to methods that are routine to one of ordinary skill in the art (Sanger, F., and Coulson, A.R., *J. Mol. Biol.* 94:444-448 (1975); Sanger, F., et al., *Proc. Natl. Acad. Sci. USA* 74:5463-30 5467 (1977)). All amino acid sequences of polypeptides encoded by DNA

-50-

5

10

15

20

25

30

molecules determined herein were predicted by conceptual translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by these approaches, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by such methods are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

Unless otherwise indicated, each "nucleotide sequence" set forth herein is presented as a sequence of deoxyribonucleotides (abbreviated A, G, C and T). However, by "nucleotide sequence" of a nucleic acid molecule or polynucleotide is intended, for a DNA molecule or polynucleotide, a sequence of deoxyribonucleotides, and for an RNA molecule or polynucleotide, the corresponding sequence of ribonucleotides (A, G, C and U), where each thymidine deoxyribonucleotide (T) in the specified deoxyribonucleotide sequence is replaced by the ribonucleotide uridine (U). Thus, the invention relates to sequences of the invention in the form of DNA or RNA molecules, or hybrid DNA/RNA molecules, and their corresponding complementary DNA, RNA, or DNA/RNA strands.

In a first such aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attB1*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attB1* nucleotide sequence having the sequence set forth in Figure 9, such as: ACAAGTTGTACAAAAAAGCAGGCT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attB1*, or mutants, fragments, variants or derivatives thereof. As one of ordinary skill will appreciate, however, certain mutations, insertions, or deletions of one or more bases in the *attB1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional

integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attB1* sequence are encompassed within the scope of the invention.

In a related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attB2*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attB2* nucleotide sequence having the sequence set forth in Figure 9, such as: ACCCAGCTTCTTGTACAAAGTGGT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attB2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attB2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attB2* sequence are encompassed within the scope of the invention.

A recombinant host cell comprising a nucleic acid molecule containing *attB1* and *attB2* sites (the vector pEXP501, also known as pCMVSport6; see Figure 48), *E. coli* DB3.1(pCMVSport6), was deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit No. NRRL B-30108. The *attB1* and *attB2* sites within the deposited nucleic acid molecule are contained in nucleic acid cassettes in association with one or more additional functional sequences as described in more detail below.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attP1*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attP1* nucleotide sequence having the sequence set forth in Figure 9, such as: TACAGGTCACTAATACCATCTAAGTAGTTGATTCATAGTGA-CTGGATATGTTGTGTTTACAGTATTATGTAGTCTGTTTTAT-GCAAAATCTAATTAAATATTGATATTATATCATTTCAGTT-TCTCGTTCAGCTTTTGACAAAGTTGGCATTATAAAAAAGCATTG-CTCATCAATTGTTGCAACGAACAGGTCACTATCAGTCAAAATAA-

AATCATTATTTG, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attP1*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attP1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attP1* sequence are encompassed within the scope of the invention.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attP2*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attP2* nucleotide sequence having the sequence set forth in Figure 9, such as: CAAATAATGATTTATTTGACTGATAGTGACCTGTTCGTTG-CAACAAATTGATAAGCAATGCTTCTTATAATGCCAACTTT-GTACAAGAAAGCTGAACGAGAACGTAAAATGATA-TAAATATCAATATATTAAATTAGATTTGCATAAAAAACAG-ACTACATAATACTGTAAAACACAACATATCCAGTCACTATGAATCAA-CTACTTAGATGGTATTAGTGACCTGTA, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attP2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attP2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attP2* sequence are encompassed within the scope of the invention.

A recombinant host cell comprising a nucleic acid molecule (the *attP* vector pDONR201, also known as pENTR21-*attPkan* or pAttPkan; see Figure 49) containing *attP1* and *attP2* sites, *E. coli* DB3.1(pAttPkan) (also called *E. coli* DB3.1(pAHKan)), was deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit No. NRRL B-30099. The *attP1* and *attP2* sites within the deposited nucleic acid molecule are contained in nucleic acid

cassettes in association with one or more additional functional sequences as described in more detail below.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attR1*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attR1* nucleotide sequence having the sequence set forth in Figure 9, such as: ACAAGTTGTACAAAAAAGCTGAACGAG-AAACGTAAAATGATATAAATATCAATATATTAAATTAGATTTCGCAT-AAAAAACAGACTACATAACTGTAAAACACAACATATCCAGTCA-CTATG, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attR1*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attR1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attR1* sequence are encompassed within the scope of the invention.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attR2*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attR2* nucleotide sequence having the sequence set forth in Figure 9, such as: GCAGGTCTGACCATAGTAGTGACTGGATAT-GTTGTGTTTACAGTATTATGTAGTCTGTTTTATGCAAATCTA-ATTTAATATATTGATATTTATATCATTACGTTCTCGTTAGCTT-TCTTGTACAAAGTGGT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attR2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attR2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attR2* sequence are encompassed within the scope of the invention.

-54-

Recombinant host cell strains containing attR1 sites apposed to cloning sites in reading frame A, reading frame B, and reading frame C, *E. coli* DB3.1(pEZC15101) (reading frame A; see Figure 64A), *E. coli* DB3.1(pEZC15102) (reading frame B; see Figure 64B), and *E. coli* DB3.1(pEZC15103) (reading frame C; see Figure 64C), and containing corresponding attR2 sites, were deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit Nos. NRRL B-30103, NRRL B-30104, and NRRL B-30105, respectively. The attR1 and attR2 sites within the deposited nucleic acid molecules are contained in nucleic acid cassettes in association with one or more additional functional sequences as described in more detail below.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attL1*, or mutants, fragments, variants and derivatives thereof. Such nucleic acid molecules may comprise an *attL1* nucleotide sequence having the sequence set forth in Figure 9, such as: CAA ATA ATG ATT TTA TTT TGA CTG ATA GTG ACC TGT TCG TTG CAA CAA ATT GAT AAG CAA TGC TTT TTT ATA ATG CCA ACT TTG TAC AAA AAA GCA GGC T, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attL1*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attL1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attL1* sequence are encompassed within the scope of the invention.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attL2*, or mutants, fragments, variants and derivatives thereof. Such nucleic acid molecules may comprise an *attL2* nucleotide sequence having the sequence set forth in Figure 9, such as: C AAA TAA TGA TTT TAT TTT GAC TGA TAG TGA CCT GTT CGT TGC AAC AAA TTG ATA AGC AAT GCT TTC TTA TAA TGC CAA

5 CTT TGT ACA AGA AAG CTG GGT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attL2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attL2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attL2* sequence are encompassed within the scope of the invention.

10 Recombinant host cell strains containing *attL1* sites apposed to cloning sites in reading frame A, reading frame B, and reading frame C, *E. coli* DB3.1(pENTR1A) (reading frame A; see Figure 10), *E. coli* DB3.1(pENTR2B) (reading frame B; see Figure 11), and *E. coli* DB3.1(pENTR3C) (reading frame C; see Figure 12), and containing corresponding *attL2* sites, were deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection 15 (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit Nos. NRRL B-30100, NRRL B-30101, and NRRL B-30102, respectively. The *attL1* and *attL2* sites within the deposited nucleic acid molecules are contained in nucleic acid cassettes in association with one or more additional functional sequences as described in more detail below.

20 Each of the recombination site sequences described herein or portions thereof, or the nucleotide sequence cassettes contained in the deposited clones, may be cloned or inserted into a vector of interest (for example, using the recombinational cloning methods described herein and/or standard restriction cloning techniques that are routine in the art) to generate, for example, Entry Vectors or Destination 25 Vectors which may be used to transfer a desired segment of a nucleic acid molecule of interest (e.g., a gene, cDNA molecule, or cDNA library) into a desired vector or into a host cell.

30 Using the information provided herein, such as the nucleotide sequences for the recombination site sequences described herein, an isolated nucleic acid molecule of the present invention encoding one or more recombination sites or portions thereof may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting material. Preferred such

methods include PCR-based cloning methods, such as reverse transcriptase-PCR (RT-PCR) using primers such as those described herein and in the Examples below. Alternatively, vectors comprising the cassettes containing the recombination site sequences described herein are available commercially from Life Technologies, Inc. (Rockville, MD).

The invention is also directed to nucleic acid molecules comprising one or more of the recombination site sequences or portions thereof and one or more additional nucleotide sequences, which may encode functional or structural sites such as one or more multiple cloning sites, one or more transcription termination sites, one or more transcriptional regulatory sequences (which may be promoters, enhancers, repressors, and the like), one or more translational signals (e.g., secretion signal sequences), one or more origins of replication, one or more fusion partner peptides (particularly glutathione S-transferase (GST), hexahistidine (His_6), and thioredoxin (Trx)), one or more selection markers or modules, one or more nucleotide sequences encoding localization signals such as nuclear localization signals or secretion signals, one or more origins of replication, one or more protease cleavage sites, one or more genes or portions of genes encoding a protein or polypeptide of interest, and one or more 5' polynucleotide extensions (particularly an extension of guanine residues ranging in length from about 1 to about 20, from about 2 to about 15, from about 3 to about 10, from about 4 to about 10, and most preferably an extension of 4 or 5 guanine residues at the 5' end of the recombination site nucleotide sequence). The one or more additional functional or structural sequences may or may not flank one or more of the recombination site sequences contained on the nucleic acid molecules of the invention.

In some nucleic acid molecules of the invention, the one or more nucleotide sequences encoding one or more additional functional or structural sites may be operably linked to the nucleotide sequence encoding the recombination site. For example, certain nucleic acid molecules of the invention may have a promoter sequence operably linked to a nucleotide sequence encoding a recombination site or portion thereof of the invention, such as a T7 promoter, a phage lambda PL

promoter, an *E. coli lac*, *trp* or *tac* promoter, and other suitable promoters which will be familiar to the skilled artisan.

Nucleic acid molecules of the present invention, which may be isolated nucleic acid molecules, may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced synthetically, or in the form of DNA-RNA hybrids. The nucleic acid molecules of the invention may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand. The nucleic acid molecules of the invention may also have a number of topologies, including linear, circular, coiled, or supercoiled.

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells, and those DNA molecules purified (partially or substantially) from a solution whether produced by recombinant DNA or synthetic chemistry techniques. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention.

The present invention further relates to mutants, fragments, variants and derivatives of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of one or more recombination sites. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism (see Lewin, B., ed., *Genes II*, John Wiley & Sons, New York (1985)). Non-naturally occurring variants may be produced using art-known mutagenesis techniques, such as those described hereinbelow.

Such variants include those produced by nucleotide substitutions, deletions or additions or portions thereof, or combinations thereof. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding

regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the encoded polypeptide(s) or portions thereof, and which also do not substantially alter the reactivities of the recombination site nucleic acid sequences in recombination reactions. Also especially preferred in this regard are conservative substitutions.

Particularly preferred mutants, fragments, variants, and derivatives of the nucleic acid molecules of the invention include, but are not limited to, insertions, deletions or substitutions of one or more nucleotide bases within the 15 bp core region (GCTTTTTTATACTAA) which is identical in all four wildtype lambda *att* sites, *attB*, *attP*, *attL* and *attR* (see U.S. Application Nos. 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed October 23, 1998, which describes the core region in further detail, and the disclosures of which are incorporated herein by reference in their entireties). Analogously, the core regions in *attB1*, *attP1*, *attL1* and *attR1* are identical to one another, as are the core regions in *attB2*, *attP2*, *attL2* and *attR2*. Particularly preferred in this regard are nucleic acid molecules comprising insertions, deletions or substitutions of one or more nucleotides within the seven bp overlap region (TTTATAC, which is defined by the cut sites for the integrase protein and is the region where strand exchange takes place) that occurs within this 15 bp core region (GCTTTTTTATACTAA). Examples of such preferred mutants, fragments, variants and derivatives according to this aspect of the invention include, but are not limited to, nucleic acid molecules in which the thymine at position 1 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; in which the thymine at position 2 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; in which the thymine at position 3 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; in which the adenine at position 4 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or thymine; in which the thymine at position 5 of the seven bp overlap region has been deleted or substituted with a

5 guanine, cytosine, or adenine; in which the adenine at position 6 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or thymine; and in which the cytosine at position 7 of the seven bp overlap region has been deleted or substituted with a guanine, thymine, or adenine; or any combination of one or more such deletions and/or substitutions within this seven bp overlap region. As described in detail in Example 21 herein, mutants of the nucleic acid molecules of the invention in which substitutions have been made within the first three positions of the seven bp overlap (TTTATAC) have been found in the present invention to strongly affect the specificity of recombination, 10 mutant nucleic acid molecules in which substitutions have been made in the last four positions (TTTATAC) only partially alter recombination specificity, and mutant nucleic acid molecules comprising nucleotide substitutions outside of the seven bp overlap, but elsewhere within the 15 bp core region, do not affect specificity of recombination but do influence the efficiency of recombination.

15 Hence, in an additional aspect, the present invention is also directed to nucleic acid molecules comprising one or more recombination site nucleotide sequences that affect recombination specificity, particularly one or more nucleotide sequences that may correspond substantially to the seven base pair overlap within the 15 bp core region, having one or more mutations that affect recombination specificity. Particularly preferred such molecules may comprise a consensus sequence (described in detail in Example 21 herein) such as NNNATAC, wherein "N" refers to any nucleotide (*i.e.*, may be A, G, T/U or C), with the proviso that if one of the first three nucleotides in the consensus sequence is a T/U, then at least one of the other two of the first three nucleotides is not a T/U.

20 25 In a related aspect, the present invention is also directed to nucleic acid molecules comprising one or more recombination site nucleotide sequences that enhance recombination efficiency, particularly one or more nucleotide sequences that may correspond substantially to the core region and having one or more mutations that enhance recombination efficiency. By sequences or mutations that "enhance recombination efficiency" is meant a sequence or mutation in a recombination site, preferably in the core region (*e.g.*, the 15 bp core region of *att* recombination sites), that results in an increase in cloning efficiency (typically

measured by determining successful cloning of a test sequence, e.g., by determining CFU/ml for a given cloning mixture) when recombining molecules comprising the mutated sequence or core region as compared to molecules that do not comprise the mutated sequence or core region (e.g., those comprising a wildtype recombination site core region sequence). More specifically, whether or not a given sequence or mutation enhances recombination efficiency may be determined using the sequence or mutation in recombinational cloning as described herein, and determining whether the sequence or mutation provides enhanced recombinational cloning efficiency when compared to a non-mutated (e.g., wildtype) sequence. Methods of determining preferred cloning efficiency-enhancing mutations for a number of recombination sites, particularly for *att* recombination sites, are described herein, for example in Examples 22-25. Examples of preferred such mutant recombination sites include but are not limited to the *attL* consensus core sequence of caactnnnnnannaagttg (wherein "n" represents any nucleotide), for example the *attL5* sequence agccctgctttattatactaagttggcatta and the *attL6* sequence agccctgcttttatattaagttggcatta; the *attB1.6* sequence ggggacaaccttgtacaaaaaagttggct; the *attB2.2* sequence ggggacaaccttgtacaagaaagctgggt; and the *attB2.10* sequence ggggacaaccttgtacaagaagttgggt. Those of skill in the art will appreciate that, in addition to the core region, other portions of the *att* site may affect the efficiency of recombination. There are five so-called arm binding sites for the integrase protein in the bacteriophage lambda *attP* site, two in *attR* (P1 and P2), and three in *attL* (P'1, P'2 and P'3). Compared to the core binding sites, the integrase protein binds to arm sites with high affinity and interacts with core and arm sites through two different domains of the protein. As with the core binding site a consensus sequence for the arm binding site consisting of C/AAGTCACTAT has been inferred from sequence comparison of the five arm binding sites and seven non-*att* sites (Ross and Landy, *Proc. Natl. Acad. Sci. USA* 79:7724-7728 (1982)). Each arm site has been mutated and tested for its effect in the excision and integration reactions (Numrych *et al.*, *Nucl. Acids Res.* 18:3953 (1990)). Hence, specific sites are utilized in each reaction in different ways, namely, the P1 and P'3

sites are essential for the integration reaction whereas the other three sites are dispensable to the integration reaction to varying degrees. Similarly, the P2, P'1 and P'2 sites are most important for the excision reaction, whereas P1 and P'3 are completely dispensable. Interestingly, when P2 is mutated the integration reaction occurs more efficiently than with the wild type attP site. Similarly, when P1 and P'3 are mutated the excision reaction occurs more efficiently. The stimulatory effect of mutating integrase arm binding sites can be explained by removing sites that compete or inhibit a specific recombination pathway or that function in a reaction that converts products back to starting substrates. In fact there is evidence for an XIS-independent LR reaction (Abremski and Gottesman, *J. Mol. Biol.* 153:67-78 (1981)). Thus, in addition to modifications in the core region of the att site, the present invention contemplates the use of att sites containing one or more modifications in the integrase arm-type binding sites. In some preferred embodiments, one or more mutations may be introduced into one or more of the P1, P'1, P2, P'2 and P'3 sites. In some preferred embodiments, multiple mutations may be introduced into one or more of these sites. Preferred such mutations include those which increase the recombination *in vitro*. For example, in some embodiments mutations may be introduced into the arm-type binding sites such that integrative recombination, corresponding to the BP reaction, is enhanced. In other embodiments, mutations may be introduced into the arm-type binding sites such that excisive recombination, corresponding to the LR reaction, is enhanced. Of course, based on the guidance contained herein, particularly in the construction and evaluation of effects of mutated recombination sites upon recombinational specificity and efficiency, analogous mutated or engineered sequences may be produced for other recombination sites described herein (including but not limited to *lox*, FRT, and the like) and used in accordance with the invention. For example, much like the mutagenesis strategy used to select core binding sites that enhance recombination efficiency, similar strategies can be employed to select changes in the arms of attP, attL and attR, and in analogous sequences in other recombination sites such as *lox*, FRT and the like, that enhance recombination efficiency. Hence, the construction and evaluation of such mutants is well within the abilities of those of ordinary skill in the art without undue experimentation.

One suitable methodology for preparing and evaluating such mutations is found in Numrych, *et al.*, (1990) *Nucleic Acids Research* 18(13): 3953-3959.

Other mutant sequences and nucleic acid molecules that may be suitable to enhance recombination efficiency will be apparent from the description herein, or 5 may be easily determined by one of ordinary skill using only routine experimentation in molecular biology in view of the description herein and information that is readily available in the art

Since the genetic code is well known in the art, it is also routine for one of ordinary skill in the art to produce degenerate variants of the nucleic acid 10 molecules described herein without undue experimentation. Hence, nucleic acid molecules comprising degenerate variants of nucleic acid sequences encoding the recombination sites described herein are also encompassed within the scope of the invention.

Further embodiments of the invention include isolated nucleic acid molecules 15 comprising a polynucleotide having a nucleotide sequence at least 50% identical, at least 60% identical, at least 70% identical, at least 75% identical, at least 80% identical, at least 85% identical, at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical to the nucleotide sequences of the seven bp overlap region within the 15 bp core region of the recombination sites 20 described herein, or the nucleotide sequences of *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* as set forth in Figure 9 (or portions thereof), or a nucleotide sequence complementary to any of these nucleotide sequences, or fragments, variants, mutants, and derivatives thereof.

By a polynucleotide having a nucleotide sequence at least, for example, 95% 25 "identical" to a reference nucleotide sequence encoding a particular recombination site or portion thereof is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations (*e.g.*, insertions, substitutions, or deletions) per each 100 nucleotides of the reference nucleotide sequence encoding the recombination site. For example, to obtain a polynucleotide having a 30 nucleotide sequence at least 95% identical to a reference *attB1* nucleotide sequence, up to 5% of the nucleotides in the *attB1* reference sequence may be

5 deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the *attB1* reference sequence may be inserted into the *attB1* reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

10 As a practical matter, whether any particular nucleic acid molecule is at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, a given recombination site nucleotide sequence or portion thereof can be determined conventionally using known computer programs such as DNAsis software (Hitachi Software, San Bruno, California) for initial sequence alignment followed by ESEE version 3.0 DNA/protein sequence software (cabot@trog.mbb.sfu.ca) for multiple sequence alignments. Alternatively, such determinations may be accomplished using the BESTFIT program (Wisconsin 15 Sequence Analysis Package, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711), which employs a local homology algorithm (Smith and Waterman, *Advances in Applied Mathematics* 2: 482-489 (1981)) to find the best segment of homology between two sequences. When using DNAsis, ESEE, BESTFIT or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number 20 of nucleotides in the reference sequence are allowed.

25

30 The present invention is directed to nucleic acid molecules at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to the *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* nucleotide sequences as set forth in Figure 9, or to the nucleotide sequence of the deposited clones, irrespective of whether they encode particular functional polypeptides. This is because even where a particular nucleic acid molecule does not encode a particular functional polypeptide, one of skill in the art would still know how to use the nucleic acid

molecule, for instance, as a hybridization probe or a polymerase chain reaction (PCR) primer.

Mutations can also be introduced into the recombination site nucleotide sequences for enhancing site specific recombination or altering the specificities of the reactants, etc. Such mutations include, but are not limited to: recombination sites without translation stop codons that allow fusion proteins to be encoded; recombination sites recognized by the same proteins but differing in base sequence such that they react largely or exclusively with their homologous partners allowing multiple reactions to be contemplated; and mutations that prevent hairpin formation of recombination sites. Which particular reactions take place can be specified by which particular partners are present in the reaction mixture.

There are well known procedures for introducing specific mutations into nucleic acid sequences. A number of these are described in Ausubel, F.M. *et al.*, *Current Protocols in Molecular Biology*, Wiley Interscience, New York (1989-1996). Mutations can be designed into oligonucleotides, which can be used to modify existing cloned sequences, or in amplification reactions. Random mutagenesis can also be employed if appropriate selection methods are available to isolate the desired mutant DNA or RNA. The presence of the desired mutations can be confirmed by sequencing the nucleic acid by well known methods.

The following non-limiting methods can be used to modify or mutate a given nucleic acid molecule encoding a particular recombination site to provide mutated sites that can be used in the present invention:

1. By recombination of two parental DNA sequences by site-specific (e.g. attL and attR to give attP) or other (e.g. homologous) recombination mechanisms where the parental DNA segments contain one or more base alterations resulting in the final mutated nucleic acid molecule;
2. By mutation or mutagenesis (site-specific, PCR, random, spontaneous, etc) directly of the desired nucleic acid molecule;
3. By mutagenesis (site-specific, PCR, random, spontaneous, etc) of parental DNA sequences, which are recombined to generate a desired nucleic acid molecule;

-65-

4. By reverse transcription of an RNA encoding the desired core sequence; and
5. By *dé novo* synthesis (chemical synthesis) of a sequence having the desired base changes, or random base changes followed by sequencing or functional analysis according to methods that are routine in the art.

The functionality of the mutant recombination sites can be demonstrated in ways that depend on the particular characteristic that is desired. For example, the lack of translation stop codons in a recombination site can be demonstrated by expressing the appropriate fusion proteins. Specificity of recombination between homologous partners can be demonstrated by introducing the appropriate molecules into *in vitro* reactions, and assaying for recombination products as described herein or known in the art. Other desired mutations in recombination sites might include the presence or absence of restriction sites, translation or transcription start signals, protein binding sites, particular coding sequences, and other known functionalities of nucleic acid base sequences. Genetic selection schemes for particular functional attributes in the recombination sites can be used according to known method steps. For example, the modification of sites to provide (from a pair of sites that do not interact) partners that do interact could be achieved by requiring deletion, via recombination between the sites, of a DNA sequence encoding a toxic substance. Similarly, selection for sites that remove translation stop sequences, the presence or absence of protein binding sites, etc., can be easily devised by those skilled in the art.

Accordingly, the present invention also provides a nucleic acid molecule, comprising at least one DNA segment having at least one, and preferably at least two, engineered recombination site nucleotide sequences of the invention flanking a selectable marker and/or a desired DNA segment, wherein at least one of said recombination site nucleotide sequences has at least one engineered mutation that enhances recombination *in vitro* in the formation of a Cointegrate DNA or a Product DNA. Such engineered mutations may be in the core sequence of the recombination site nucleotide sequence of the invention; see U.S. Application Nos. 08/486,139, filed June 7, 1995, 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed

October 23, 1998, the disclosures of which are all incorporated herein by reference in their entireties.

While in the preferred embodiment the recombination sites differ in sequence and do not interact with each other, it is recognized that sites comprising the same sequence, which may interact with each other, can be manipulated or engineered to inhibit recombination with each other. Such conceptions are considered and incorporated herein. For example, a protein binding site (*e.g.*, an antibody-binding site, a histone-binding site, an enzyme-binding site, or a binding site for any nucleic acid molecule-binding protein) can be engineered adjacent to one of the sites. In the presence of the protein that recognizes the engineered site, the recombinase fails to access the site and another recombination site in the nucleic acid molecule is therefore used preferentially. In the cointegrate this site can no longer react since it has been changed, *e.g.*, from attB to attL. During or upon resolution of the cointegrate, the protein can be inactivated (*e.g.*, by antibody, heat or a change of buffer) and the second site can undergo recombination.

The nucleic acid molecules of the invention can have at least one mutation that confers at least one enhancement of said recombination, said enhancement selected from the group consisting of substantially (i) favoring integration; (ii) favoring recombination; (ii) relieving the requirement for host factors; (iii) increasing the efficiency of said Cointegrate DNA or Product DNA formation; (iv) increasing the specificity of said Cointegrate DNA or Product DNA formation; and (v) adding or deleting protein binding sites.

In other embodiments, the nucleic acid molecules of the invention may be PCR primer molecules, which comprise one or more of the recombination site sequences described herein or portions thereof, particularly those shown in Figure 9 (or sequences complementary to those shown in Figure 9), or mutants, fragments, variants or derivatives thereof, attached at the 3' end to a target-specific template sequence which specifically interacts with a target nucleic acid molecule which is to be amplified. Primer molecules according to this aspect of the invention may further comprise one or more, (*e.g.*, 1, 2, 3, 4, 5, 10, 20, 25, 50, 100, 500, 1000, or more) additional bases at their 5' ends, and preferably comprise one or more (particularly four or five) additional bases, which are preferably

guanines, at their 5' ends, to increase the efficiency of the amplification products incorporating the primer molecules in the recombinational cloning system of the invention. Such nucleic acid molecules and primers are described in detail in the examples herein, particularly in Examples 22-25.

5 Certain primers of the invention may comprise one or more nucleotide deletions in the *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* sequences as set forth in Figure 9. In one such aspect, for example, *attB2* primers may be constructed in which one or more of the first four nucleotides at the 5' end of the *attB2* sequence shown in Figure 9 have been deleted. Primers according to this aspect of the invention may therefore have the sequence:

10 (*attB2(-1)*): CCCAGCTTCTTGTACAAAGTGGTnnnnnnnnnnnn . . . n

(*attB2(-2)*): CCAGCTTCTTGTACAAAGTGGTnnnnnnnnnnnn . . . n

(*attB2(-3)*): CAGCTTCTTGTACAAAGTGGTnnnnnnnnnnnn . . . n

(*attB2(-4)*): AGCTTCTTGTACAAAGTGGTnnnnnnnnnnnn . . . n,

15 wherein "nnnnnnnnnnnn . . . n" at the 3' end of the primer represents a target-specific sequence of any length, for example from one base up to all of the bases of a target nucleic acid molecule (e.g., a gene) or a portion thereof, the sequence and length which will depend upon the identity of the target nucleic acid molecule which is to be amplified.

20 The primer nucleic acid molecules according to this aspect of the invention may be produced synthetically by attaching the recombination site sequences depicted in Figure 9, or portions thereof, to the 5' end of a standard PCR target-specific primer according to methods that are well-known in the art. Alternatively, additional primer nucleic acid molecules of the invention may be produced synthetically by adding one or more nucleotide bases, which preferably correspond to one or more, preferably five or more, and more preferably six or more, contiguous nucleotides of the *att* nucleotide sequences described herein (see, e.g., Example 20 herein; see also U.S. Application Nos. 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed October 23, 1998, the disclosures of which are all incorporated herein by reference in their entireties), to the 5' end of a standard PCR target-specific primer according to methods that are well-known in the art, to provide

-68-

primers having the specific nucleotide sequences described herein. As noted above, primer nucleic acid molecules according to this aspect of the invention may also optionally comprise one, two, three, four, five, or more additional nucleotide bases at their 5' ends, and preferably will comprise four or five guanines at their 5' ends. In one particularly preferred such aspect, the primer nucleic acid molecules of the invention may comprise one or more, preferably five or more, more preferably six or more, still more preferably 6-18 or 6-25, and most preferably 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25, contiguous nucleotides or bp of the *att*B1 or *att*B2 nucleotide sequences depicted in Figure 9 (or nucleotides complementary thereto), linked to the 5' end of a target-specific (e.g., a gene-specific) primer molecule. Primer nucleic acid molecules according to this aspect of the invention include, but are not limited to, *att*B1- and *att*B2-derived primer nucleic acid molecules having the following nucleotide sequences:

15 ACAAGTTGTACAAAAAAGCAGGCT-nnnnnnnnnnnnn . . . n
 ACCACTTGTACAAGAAAGCTGGGT-nnnnnnnnnnnnn . . . n
 TGTACAAAAAAGCAGGCT-nnnnnnnnnnnnn . . . n
 TGTACAAGAAAGCTGGGT-nnnnnnnnnnnnn . . . n
 ACAAAAAAAGCAGGCT-nnnnnnnnnnnnn . . . n
20 ACAAGAAAGCTGGGT-nnnnnnnnnnnnn . . . n
 AAAAAGCAGGCT-nnnnnnnnnnnnn . . . n
 AGAAAGCTGGGT-nnnnnnnnnnnnn . . . n
 AAAAGCAGGCT-nnnnnnnnnnnnn . . . n
 GAAAGCTGGGT-nnnnnnnnnnnnn . . . n
25 AAAGCAGGCT-nnnnnnnnnnnnn . . . n
 AAAGCTGGGT-nnnnnnnnnnnnn . . . n
 AAGCAGGCT-nnnnnnnnnnnnn . . . n
 AAGCTGGGT-nnnnnnnnnnnnn . . . n
 AGCAGGCT-nnnnnnnnnnnnn . . . n
30 AGCTGGGT-nnnnnnnnnnnnn . . . n
 GCAGGCT-nnnnnnnnnnnnn . . . n
 GCTGGGT-nnnnnnnnnnnnn . . . n

CAGGCT-nnnnnnnnnnnnn . . . n

CTGGGT-nnnnnnnnnnnnn . . . n,

wherein "nnnnnnnnnnnn . . . n" at the 3' end of the primer represents a target-specific sequence of any length, for example from one base up to all of the bases of a target nucleic acid molecule (e.g., a gene) or a portion thereof, the sequence and length which will depend upon the identity of the target nucleic acid molecule which is to be amplified.

Of course, it will be apparent to one of ordinary skill from the teachings contained herein that additional primer nucleic acid molecules analogous to those specifically described herein may be produced using one or more, preferably five or more, more preferably six or more, still more preferably ten or more, 15 or more, 20 or more, 25 or more, 30 or more, etc. (through to and including all) of the contiguous nucleotides or bp of the *att*P1, *att*P2, *att*L1, *att*L2, *att*R1 or *att*R2 nucleotide sequences depicted in Figure 9 (or nucleotides complementary thereto), linked to the 5' end of a target-specific (e.g., a gene-specific) primer molecule. As noted above, such primer nucleic acid molecules may optionally further comprise one, two, three, four, five, or more additional nucleotide bases at their 5' ends, and preferably will comprise four guanines at their 5' ends. Other primer molecules comprising the *att*B1, *att*B2, *att*P1, *att*P2, *att*L1, *att*L2, *att*R1 and *att*R2 sequences depicted in Figure 9, or portions thereof, may be made by one of ordinary skill without resorting to undue experimentation in accordance with the guidance provided herein.

The primers of the invention described herein are useful in producing PCR fragments having a nucleic acid molecule of interest flanked at each end by a recombination site sequence (as described in detail below in Example 9), for use in cloning of PCR-amplified DNA fragments using the recombination system of the invention (as described in detail below in Examples 8, 19 and 21-25).

Vectors

The invention also relates to vectors comprising one or more of the nucleic acid molecules of the invention, as described herein. In accordance with the invention, any vector may be used to construct the vectors of the invention. In

particular, vectors known in the art and those commercially available (and variants or derivatives thereof) may in accordance with the invention be engineered to include one or more nucleic acid molecules encoding one or more recombination sites (or portions thereof), or mutants, fragments, or derivatives thereof, for use in the methods of the invention. Such vectors may be obtained from, for example, 5 Vector Laboratories Inc., InVitrogen, Promega, Novagen, New England Biolabs, Clontech, Roche, Pharmacia, EpiCenter, OriGenes Technologies Inc., Stratagene, Perkin Elmer, Pharmingen, Life Technologies, Inc., and Research Genetics. Such vectors may then for example be used for cloning or subcloning nucleic acid 10 molecules of interest. General classes of vectors of particular interest include prokaryotic and/or eukaryotic cloning vectors, Expression Vectors, fusion vectors, two-hybrid or reverse two-hybrid vectors, shuttle vectors for use in different hosts, mutagenesis vectors, transcription vectors, vectors for receiving large 15 inserts and the like.

Other vectors of interest include viral origin vectors (M13 vectors, bacterial 20 phage λ vectors, bacteriophage P1 vectors, adenovirus vectors, herpesvirus vectors, retrovirus vectors, phage display vectors, combinatorial library vectors), high, low, and adjustable copy number vectors, vectors which have compatible replicons for use in combination in a single host (pACYC184 and pBR322) and eukaryotic episomal replication vectors (pCDM8).

Particular vectors of interest include prokaryotic Expression Vectors such as pcDNA II, pSL301, pSE280, pSE380, pSE420, pTrcHisA, B, and C, pRSET A, B, and C (Invitrogen, Inc.), pGEMEX-1, and pGEMEX-2 (Promega, Inc.), the pET vectors (Novagen, Inc.), pTrc99A, pKK223-3, the pGEX vectors, pEZ218, pRIT2T, and pMC1871 (Pharmacia, Inc.), pKK233-2 and pKK388-1 (Clontech, Inc.), and pProEx-HT (Life Technologies, Inc.) and variants and derivatives 25 thereof. Destination Vectors can also be made from eukaryotic Expression Vectors such as pFastBac, pFastBac HT, pFastBac DUAL, pSFV, and pTet-Splice (Life Technologies, Inc.), pEUK-C1, pPUR, pMAM, pMAMneo, pBI101, pBI121, pDR2, pCMVEBNA, and pYACneo (Clontech), pSVK3, pSVL, pMSG, pCH110, and pKK232-8 (Pharmacia, Inc.), p3'SS, pXT1, pSG5, pPbac, pMbac, 30 pMC1neo, and pOG44 (Stratagene, Inc.), and pYES2, pAC360, pBlueBacHis A,

-71-

B, and C, pVL1392, pBsueBacIII, pCDM8, pcDNA1, pZeoSV, pcDNA3 pREP4, pCEP4, and pEBVHis (Invitrogen, Inc.) and variants or derivatives thereof.

Other vectors of particular interest include pUC18, pUC19, pBlueScript, pSPORT, cosmids, phagemids, YACs (yeast artificial chromosomes), BACs (bacterial artificial chromosomes), MACs (mammalian artificial chromosomes), pQE70, pQE60, pQE9 (Qiagen), pBS vectors, PhageScript vectors, BlueScript vectors, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene), pcDNA3 (InVitrogen), pGEX, pTrsfus, pTrc99A, pET-5, pET-9, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia), pSPORT1, pSPORT2, pCMVSPORT2.0 and pSV-SHORT1 (Life Technologies, Inc.) and variants or derivatives thereof.

Additional vectors of interest include pTrxFus, pThioHis, pLEX, pTrcHis, pTrcHis2, pRSET, pBlueBacHis2, pcDNA3.1/His, pcDNA3.1(-)/Myc-His, pSecTag, pEBVHis, pPIC9K, pPIC3.5K, pAO815, pPICZ, pPICZ α , pGAPZ, pGAPZ α , pBlueBac4.5, pBlueBacHis2, pMelBac, pSinRep5, pSinHis, pIND, pIND(SP1), pVgRXR, pcDNA2.1, pYES2, pZErO1.1, pZErO-2.1, pCR-Blunt, pSE280, pSE380, pSE420, pVL1392, pVL1393, pCDM8, pcDNA1.1, pcDNA1.1/Amp, pcDNA3.1, pcDNA3.1/Zeo, pSe, SV2, pRc/CMV2, pRc/RSV, pREP4, pREP7, pREP8, pREP9, pREP10, pCEP4, pEBVHis, pCR3.1, pCR2.1, pCR3.1-Uni, and pCRBac from Invitrogen; λ ExCell, λ gt11, pTrc99A, pKK223-3, pGEX-1 λ T, pGEX-2T, pGEX-2TK, pGEX-4T-1, pGEX-4T-2, pGEX-4T-3, pGEX-3X, pGEX-5X-1, pGEX-5X-2, pGEX-5X-3, pEZZ18, pRIT2T, pMC1871, pSVK3, pSVL, pMSG, pCH110, pKK232-8, pSL1180, pNEO, and pUC4K from Pharmacia; pSCREEN-1b(+), pT7Blue(R), pT7Blue-2, pCITE-4abc(+), pOCUS-2, pTAG, pET-32 LIC, pET-30 LIC, pBAC-2cp LIC, pBACgus-2cp LIC, pT7Blue-2 LIC, pT7Blue-2, λ SCREEN-1, λ BlueSTAR, pET-3abcd, pET-7abc, pET9abcd, pET11abcd, pET12abc, pET-14b, pET-15b, pET-16b, pET-17b-pET-17xb, pET-19b, pET-20b(+), pET-21abcd(+), pET-22b(+), pET-23abcd(+), pET-24abcd(+), pET-25b(+), pET-26b(+), pET-27b(+), pET-28abc(+), pET-29abc(+), pET-30abc(+), pET-31b(+), pET-32abc(+), pET-33b(+), pBAC-1, pBACgus-1, pBAC4x-1, pBACgus4x-1, pBAC-3cp, pBACgus-2cp, pBACsurf-1, plg, Signal plg, pYX, Selecta Vecta-Neo, Selecta Vecta - Hyg, and Selecta Vecta - Gpt from Novagen; pLexA, pB42AD, pGBT9, pAS2-1,

5 pGAD424, pACT2, pGAD GL, pGAD GH, pGAD10, pGilda, pEZM3, pEGFP,
pEGFP-1, pEGFP-N, pEGFP-C, pEBFP, pGFPuv, pGFP, p6xHis-GFP, pSEAP2-
Basic, pSEAP2-Contral, pSEAP2-Promoter, pSEAP2-Enhancer, p β gal-Basic,
p β gal-Control, p β gal-Promoter, p β gal-Enhancer, pCMV β , pTet-Off, pTet-On,
pTK-Hyg, pRetro-Off, pRetro-On, pIRES1neo, pIRES1hyg, pLXSN, pLNCX,
pLAPSN, pMAMneo, pMAMneo-CAT, pMAMneo-LUC, pPUR, pSV2neo,
- pYEX 4T-1/2/3, pYEX-S1, pBacPAK-His, pBacPAK8/9, pAcUW31, BacPAK6,
- pTriplEx, λ gt10, λ gt11, pWE15, and λ TriplEx from Clontech; Lambda ZAP II,
pBK-CMV, pBK-RSV, pBluescript II KS +/-, pBluescript II SK +/-, pAD-GAL4,
10 pBD-GAL4 Cam, pSurfscript, Lambda FIX II, Lambda DASH, Lambda EMBL3,
Lambda EMBL4, SuperCos, pCR-Script Amp, pCR-Script Cam, pCR-Script
Direct, pBS +/-, pBC KS +/-, pBC SK +/-, Phagescript, pCAL-n-EK, pCAL-n,
pCAL-c, pCAL-kc, pET-3abcd, pET-11abcd, pSPUTK, pESP-1, pCMVLaci,
15 pOPRSVI/MCS, pOPI3 CAT, pXT1, pSG5, pPbac, pMbac, pMC1neo, pMC1neo
Poly A, pOG44, pOG45, pFRT β GAL, pNEO β GAL, pRS403, pRS404, pRS405,
pRS406, pRS413, pRS414, pRS415, and pRS416 from Stratagene.

20 Two-hybrid and reverse two-hybrid vectors of particular interest include
pPC86, pDBLeu, pDBTrp, pPC97, p2.5, pGAD1-3, pGAD10, pACt, pACT2,
pGADGL, pGADGH, pAS2-1, pGAD424, pGBT8, pGBT9, pGAD-GAL4,
pLexA, pBD-GAL4, pHISi, pHISi-1, placZi, pB42AD, pDG202, pJK202,
pJG4-5, pNLexA, pYESTrp and variants or derivatives thereof.

25 Yeast Expression Vectors of particular interest include pESP-1, pESP-2,
pESC-His, pESC-Trp, pESC-URA, pESC-Leu (Stratagene), pRS401, pRS402,
pRS411, pRS412, pRS421, pRS422, and variants or derivatives thereof.

30 According to the invention, the vectors comprising one or more nucleic acid
molecules encoding one or more recombination sites, or mutants, variants,
fragments, or derivatives thereof, may be produced by one of ordinary skill in the
art without resorting to undue experimentation using standard molecular biology
methods. For example, the vectors of the invention may be produced by
introducing one or more of the nucleic acid molecules encoding one or more
recombination sites (or mutants, fragments, variants or derivatives thereof) into
one or more of the vectors described herein, according to the methods described,

for example, in Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1982). In a related aspect of the invention, the vectors may be engineered to contain, in addition to one or more nucleic acid molecules encoding one or more recombination sites (or portions thereof), one or more additional physical or functional nucleotide sequences, such as those encoding one or more multiple cloning sites, one or more transcription termination sites, one or more transcriptional regulatory sequences (e.g., one or more promoters, enhancers, or repressors), one or more selection markers or modules, one or more genes or portions of genes encoding a protein or polypeptide of interest, one or more translational signal sequences, one or more nucleotide sequences encoding a fusion partner protein or peptide (e.g., GST, His₆ or thioredoxin), one or more origins of replication, and one or more 5' or 3' polynucleotide tails (particularly a poly-G tail). According to this aspect of the invention, the one or more recombination site nucleotide sequences (or portions thereof) may optionally be operably linked to the one or more additional physical or functional nucleotide sequences described herein.

Preferred vectors according to this aspect of the invention include, but are not limited to: pENTR1A (Figures 10A and 10B), pENTR2B (Figures 11A and 11B), pENTR3C (Figures 12A and 12B), pENTR4 (Figures 13A and 13B), pENTR5 (Figures 14A and 14B), pENTR6 (Figures 15A and 15B), pENTR7 (Figures 16A and 16B), pENTR8 (Figures 17A and 17B), pENTR9 (Figures 18A and 18B), pENTR10 (Figures 19A and 19B), pENTR11 (Figures 20A and 20B), pDEST1 (Figures 21A-D), pDEST2 (Figure 22A-D), pDEST3 (Figure 23A-D), pDEST4 (Figure 24A-D), pDEST5 (Figure 25A-D), pDEST6 (Figure 26A-D), pDEST7 (Figure 27A-C), pDEST8 (Figure 28A-D), pDEST9 (Figure 29A-E), pDEST10 (Figure 30A-D), pDEST11 (Figure 31A-D), pDEST12.2 (also known as pDEST12) (Figure 32A-D), pDEST13 (Figure 33A-C), pDEST14 (Figure 34A-D), pDEST15 (Figure 35A-D), pDEST16 (Figure 36A-D), pDEST17 (Figure 37A-D), pDEST18 (Figure 38A-D), pDEST19 (Figure 39A-D), pDEST20 (Figure 40A-D), pDEST21 (Figure 41A-E), pDEST22 (Figure 42A-D), pDEST23 (Figure 43A-D), pDEST24 (Figure 44A-D), pDEST25 (Figure 45A-D), pDEST26 (Figure 46A-D), pDEST27 (Figure 47A-D), pEXP501 (also known

-74-

as pCMVSPORT6) (Figure 48A-B), pDONR201 (also known as pENTR21 attP vector or pAttPkan Donor Vector) (Figure 49), pDONR202 (Figure 50), pDONR203 (also known as pEZ15812) (Figure 51), pDONR204 (Figure 52), pDONR205 (Figure 53), pDONR206 (also known as pENTR22 attP vector or pAttPgen Donor Vector) (Figure 54), pMAB58 (Figure 87), pMAB62 (Figure 88), pDEST28 (Figure 90), pDEST29 (Figure 91), pDEST30 (Figure 92), pDEST31 (Figure 93), pDEST32 (Figure 94), pDEST33 (Figure 95), pDEST34 (Figure 96), pDONR207 (Figure 97), pMAB85 (Figure 98), pMAB86 (Figure 99), and fragments, mutants, variants, and derivatives thereof. However, it will be understood by one of ordinary skill that the present invention also encompasses other vectors not specifically designated herein, which comprise one or more of the isolated nucleic acid molecules of the invention encoding one or more recombination sites or portions thereof (or mutants, fragments, variants or derivatives thereof), and which may further comprise one or more additional physical or functional nucleotide sequences described herein which may optionally be operably linked to the one or more nucleic acid molecules encoding one or more recombination sites or portions thereof. Such additional vectors may be produced by one of ordinary skill according to the guidance provided in the present specification.

20

Polymerases

25

Preferred polypeptides having reverse transcriptase activity (*i.e.*, those polypeptides able to catalyze the synthesis of a DNA molecule from an RNA template) for use in accordance with the present invention include, but are not limited to Moloney Murine Leukemia Virus (M-MLV) reverse transcriptase, Rous Sarcoma Virus (RSV) reverse transcriptase, Avian Myeloblastosis Virus (AMV) reverse transcriptase, Rous Associated Virus (RAV) reverse transcriptase, Myeloblastosis Associated Virus (MAV) reverse transcriptase, Human Immunodeficiency Virus (HIV) reverse transcriptase, retroviral reverse transcriptase, retrotransposon reverse transcriptase, hepatitis B reverse transcriptase, cauliflower mosaic virus reverse transcriptase and bacterial reverse transcriptase. Particularly preferred are those polypeptides having reverse

30

-75-

transcriptase activity that are also substantially reduced in RNase H activity (*i.e.*, "RNase H" polypeptides). By a polypeptide that is "substantially reduced in RNase H activity" is meant that the polypeptide has less than about 20%, more preferably less than about 15%, 10% or 5%, and most preferably less than about 2%, of the RNase H activity of a wildtype or RNase H⁻ enzyme such as wildtype M-MLV reverse transcriptase. The RNase H activity may be determined by a variety of assays, such as those described, for example, in U.S. Patent No. 5,244,797, in Kotewicz, M.L. *et al.*, *Nucl. Acids Res.* 16:265 (1988) and in Gerard, G.F., *et al.*, *FOCUS* 14(5):91 (1992), the disclosures of all of which are fully incorporated herein by reference. Suitable RNase H⁻ polypeptides for use in the present invention include, but are not limited to, M-MLV H⁻ reverse transcriptase, RSV H⁻ reverse transcriptase, AMV H⁻ reverse transcriptase, RAV H⁻ reverse transcriptase, MAV H⁻ reverse transcriptase, HIV H⁻ reverse transcriptase, THERMOSCRIPT™ reverse transcriptase and THERMOSCRIPT™ II reverse transcriptase, and SUPERSCRIPT™ I reverse transcriptase and SUPERSCRIPT™ II reverse transcriptase, which are obtainable, for example, from Life Technologies, Inc. (Rockville, Maryland). See generally published PCT application WO 98/47912.

Other polypeptides having nucleic acid polymerase activity suitable for use in the present methods include thermophilic DNA polymerases such as DNA polymerase I, DNA polymerase III, Klenow fragment, T7 polymerase, and T5 polymerase, and thermostable DNA polymerases including, but not limited to, *Thermus thermophilus* (*Tth*) DNA polymerase, *Thermus aquaticus* (*Taq*) DNA polymerase, *Thermotoga neopolitana* (*Tne*) DNA polymerase, *Thermotoga maritima* (*Tma*) DNA polymerase, *Thermococcus litoralis* (*Tli* or VENT®) DNA polymerase, *Pyrococcus furiosus* (*Pfu*) DNA polymerase, *Pyrococcus* species GB-D (or DEEPVENT®) DNA polymerase, *Pyrococcus woosii* (*Pwo*) DNA polymerase, *Bacillus stearothermophilus* (*Bst*) DNA polymerase, *Sulfolobus acidocaldarius* (*Sac*) DNA polymerase, *Thermoplasma acidophilum* (*Tac*) DNA polymerase, *Thermus flavus* (*Tfl/Tub*) DNA polymerase, *Thermus ruber* (*Tru*) DNA polymerase, *Thermus brockianus* (DYNAZYME®) DNA polymerase, *Methanobacterium thermoautotrophicum* (*Mth*) DNA polymerase, and mutants,

variants and derivatives thereof. Such polypeptides are available commercially, for example from Life Technologies, Inc. (Rockville, MD), New Englan BioLabs (Beverly, MA), and Sigma/Aldrich (St. Louis, MO).

5

Host Cells

10

The invention also relates to host cells comprising one or more of the nucleic acid molecules or vectors of the invention, particularly those nucleic acid molecules and vectors described in detail herein. Representative host cells that may be used according to this aspect of the invention include, but are not limited to, bacterial cells, yeast cells, plant cells and animal cells. Preferred bacterial host cells include *Escherichia* spp. cells (particularly *E. coli* cells and most particularly *E. coli* strains DH10B, Stbl2, DH5 α , DB3, DB3.1 (preferably *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells; Life Technologies, Inc., Rockville, MD), DB4 and DB5; see U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, the disclosure of which is incorporated by reference herein in its entirety), *Bacillus* spp. cells (particularly *B. subtilis* and *B. megaterium* cells), *Streptomyces* spp. cells, *Erwinia* spp. cells, *Klebsiella* spp. cells, *Serratia* spp. cells (particularly *S. marcessans* cells), *Pseudomonas* spp. cells (particularly *P. aeruginosa* cells), and *Salmonella* spp. cells (particularly *S. typhimurium* and *S. typhi* cells). Preferred animal host cells include insect cells (most particularly *Drosophila melanogaster* cells, *Spodoptera frugiperda* Sf9 and Sf21 cells and *Trichoplusia* High-Five cells), nematode cells (particularly *C. elegans* cells), avian cells, amphibian cells (particularly *Xenopus laevis* cells), reptilian cells, and mammalian cells (most particularly CHO, COS, VERO, BHK and human cells). Preferred yeast host cells include *Saccharomyces cerevisiae* cells and *Pichia pastoris* cells. These and other suitable host cells are available commercially, for example from Life Technologies, Inc. (Rockville, Maryland), American Type Culture Collection (Manassas, Virginia), and Agricultural Research Culture Collection (NRRL; Peoria, Illinois).

25

30

Methods for introducing the nucleic acid molecules and/or vectors of the invention into the host cells described herein, to produce host cells comprising one or more of the nucleic acid molecules and/or vectors of the invention, will be

familiar to those of ordinary skill in the art. For instance, the nucleic acid molecules and/or vectors of the invention may be introduced into host cells using well known techniques of infection, transduction, transfection, and transformation. The nucleic acid molecules and/or vectors of the invention may be introduced alone or in conjunction with other the nucleic acid molecules and/or vectors. Alternatively, the nucleic acid molecules and/or vectors of the invention may be introduced into host cells as a precipitate, such as a calcium phosphate precipitate, or in a complex with a lipid. Electroporation also may be used to introduce the nucleic acid molecules and/or vectors of the invention into a host. Likewise, such molecules may be introduced into chemically competent cells such as *E. coli*. If the vector is a virus, it may be packaged *in vitro* or introduced into a packaging cell and the packaged virus may be transduced into cells. Hence, a wide variety of techniques suitable for introducing the nucleic acid molecules and/or vectors of the invention into cells in accordance with this aspect of the invention are well known and routine to those of skill in the art. Such techniques are reviewed at length, for example, in Sambrook, J., *et al.*, *Molecular Cloning, a Laboratory Manual*, 2nd Ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 16.30-16.55 (1989), Watson, J.D., *et al.*, *Recombinant DNA*, 2nd Ed., New York: W.H. Freeman and Co., pp. 213-234 (1992), and Winnacker, E.-L., *From Genes to Clones*, New York: VCH Publishers (1987), which are illustrative of the many laboratory manuals that detail these techniques and which are incorporated by reference herein in their entireties for their relevant disclosures.

Polypeptides

In another aspect, the invention relates to polypeptides encoded by the nucleic acid molecules of the invention (including polypeptides and amino acid sequences encoded by all possible reading frames of the nucleic acid molecules of the invention), and to methods of producing such polypeptides. Polypeptides of the present invention include purified or isolated natural products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, insect, mammalian, avian and higher plant cells.

-78-

The polypeptides of the invention may be produced by synthetic organic chemistry, and are preferably produced by standard recombinant methods, employing one or more of the host cells of the invention comprising the vectors or isolated nucleic acid molecules of the invention. According to the invention, 5 polypeptides are produced by cultivating the host cells of the invention (which comprise one or more of the nucleic acid molecules of the invention, preferably contained within an Expression Vector) under conditions favoring the expression of the nucleotide sequence contained on the nucleic acid molecule of the invention, such that the polypeptide encoded by the nucleic acid molecule of the invention is produced by the host cell. As used herein, "conditions favoring the expression 10 of the nucleotide sequence" or "conditions favoring the production of a polypeptide" include optimal physical (e.g., temperature, humidity, etc.) and nutritional (e.g., culture medium, ionic) conditions required for production of a recombinant polypeptide by a given host cell. Such optimal conditions for a 15 variety of host cells, including prokaryotic (bacterial), mammalian, insect, yeast, and plant cells will be familiar to one of ordinary skill in the art, and may be found, for example, in Sambrook, J., et al., *Molecular Cloning, A Laboratory Manual*, 2nd Ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, (1989), Watson, J.D., et al., *Recombinant DNA*, 2nd Ed., New York: W.H. Freeman and Co., and Winnacker, E.-L., *From Genes to Clones*, New York: VCH Publishers 20 (1987).

In some aspects, it may be desirable to isolate or purify the polypeptides of the invention (e.g., for production of antibodies as described below), resulting in the production of the polypeptides of the invention in isolated form. The polypeptides 25 of the invention can be recovered and purified from recombinant cell cultures by well-known methods of protein purification that are routine in the art, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. For example, His6 or GST fusion tags on polypeptides made by the methods of the invention may be isolated using 30 appropriate affinity chromatography matrices which bind polypeptides bearing

His₆ or GST tags, as will be familiar to one of ordinary skill in the art. Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

Isolated polypeptides of the invention include those comprising the amino acid sequences encoded by one or more of the reading frames of the polynucleotides comprising one or more of the recombination site-encoding nucleic acid molecules of the invention, including those encoding *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* having the nucleotide sequences set forth in Figure 9 (or nucleotide sequences complementary thereto), or fragments, variants, mutants and derivatives thereof; the complete amino acid sequences encoded by the polynucleotides contained in the deposited clones described herein; the amino acid sequences encoded by polynucleotides which hybridize under stringent hybridization conditions to polynucleotides having the nucleotide sequences encoding the recombination site sequences of the invention as set forth in Figure 9 (or a nucleotide sequence complementary thereto); or a peptide or polypeptide comprising a portion or a fragment of the above polypeptides. The invention also relates to additional polypeptides having one or more additional amino acids linked (typically by peptidyl bonds to form a nascent polypeptide) to the polypeptides encoded by the recombination site nucleotide sequences or the deposited clones. Such additional amino acid residues may comprise one or more functional peptide sequences, for example one or more fusion partner peptides (*e.g.*, GST, His₆, Trx, etc.) and the like.

As used herein, the terms "protein," "peptide," "oligopeptide" and "polypeptide" are considered synonymous (as is commonly recognized) and each term can be used interchangeably as the context requires to indicate a chain of two or more amino acids, preferably five or more amino acids, or more preferably ten

-80-

or more amino acids, coupled by (a) peptidyl linkage(s), unless otherwise defined in the specific contexts below. As is commonly recognized in the art, all polypeptide formulas or sequences herein are written from left to right and in the direction from amino terminus to carboxy terminus.

5 It will be recognized by those of ordinary skill in the art that some amino acid sequences of the polypeptides of the invention can be varied without significant effect on the structure or function of the polypeptides. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine structure and activity. In general, it is
10 possible to replace residues which form the tertiary structure, provided that residues performing a similar function are used. In other instances, the type of residue may be completely unimportant if the alteration occurs at a non-critical region of the polypeptide.

15 Thus, the invention further includes variants of the polypeptides of the invention, including allelic variants, which show substantial structural homology to the polypeptides described herein, or which include specific regions of these polypeptides such as the portions discussed below. Such mutants may include deletions, insertions, inversions, repeats, and type substitutions (for example, substituting one hydrophilic residue for another, but not strongly hydrophilic for
20 strongly hydrophobic as a rule). Small changes or such "neutral" or "conservative" amino acid substitutions will generally have little effect on activity.

25 Typical conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxylated residues Ser and Thr; exchange of the acidic residues Asp and Glu; substitution between the amidated residues Asn and Gln; exchange of the basic residues Lys and Arg; and replacements among the aromatic residues Phe and Tyr.

30 Thus, the fragment, derivative or analog of the polypeptides of the invention, such as those comprising peptides encoded by the recombination site nucleotide sequences described herein, may be (i) one in which one or more of the amino acid residues are substituted with a conservative or non-conservative amino acid residue (preferably a conservative amino acid residue), and such substituted amino acid residue may be encoded by the genetic code or may be an amino acid (e.g.,

-81-

desmosine, citrulline, ornithine, etc.) that is not encoded by the genetic code; (ii) one in which one or more of the amino acid residues includes a substituent group (e.g., a phosphate, hydroxyl, sulfate or other group) in addition to the normal "R" group of the amino acid; (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which additional amino acids are fused to the mature polypeptide, such as an immunoglobulin Fc region peptide, a leader or secretory sequence, a sequence which is employed for purification of the mature polypeptide (such as GST) or a proprotein sequence. Such fragments, derivatives and analogs are intended to be encompassed by the present invention, and are within the scope of those skilled in the art from the teachings herein and the state of the art at the time of invention.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. Recombinantly produced versions of the polypeptides of the invention can be substantially purified by the one-step method described in Smith and Johnson, *Gene* 67:31-40 (1988). As used herein, the term "substantially purified" means a preparation of an individual polypeptide of the invention wherein at least 50%, preferably at least 60%, 70%, or 75% and more preferably at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% (by mass) of contaminating proteins (*i.e.*, those that are not the individual polypeptides described herein or fragments, variants, mutants or derivatives thereof) have been removed from the preparation.

The polypeptides of the present invention include those which are at least about 50% identical, at least 60% identical, at least 65% identical, more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identical, to the polypeptides described herein. For example, preferred *attB1*-containing polypeptides of the invention include those that are at least about 50% identical, at least 60% identical, at least 65% identical, more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identical,

to the polypeptide(s) encoded by the three reading frames of a polynucleotide comprising a nucleotide sequence of *attB1* having a nucleic acid sequence as set forth in Figure 9 (or a nucleic acid sequence complementary thereto), to a polypeptide encoded by a polynucleotide contained in the deposited cDNA clones described herein, or to a polypeptide encoded by a polynucleotide hybridizing under stringent conditions to a polynucleotide comprising a nucleotide sequence of *attB1* having a nucleic acid sequence as set forth in Figure 9 (or a nucleic acid sequence complementary thereto). Analogous polypeptides may be prepared that are at least about 65% identical, more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identical, to the *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* polypeptides of the invention as depicted in Figure 9. The present polypeptides also include portions or fragments of the above-described polypeptides with at least 5, 10, 15, 20, or 25 amino acids.

By a polypeptide having an amino acid sequence at least, for example, 65% "identical" to a reference amino acid sequence of a given polypeptide of the invention is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to 35 amino acid alterations per each 100 amino acids of the reference amino acid sequence of a given polypeptide of the invention. In other words, to obtain a polypeptide having an amino acid sequence at least 65% identical to a reference amino acid sequence, up to 35% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 35% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino (N-) or carboxy (C-) terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence. As a practical matter, whether a given amino acid sequence is, for example, at least 65% identical to the amino acid sequence of a given polypeptide of the invention can be determined

-83-

conventionally using known computer programs such as those described above for nucleic acid sequence identity determinations, or more preferably using the CLUSTAL W program (Thompson, J.D., et al., *Nucleic Acids Res.* 22:4673-4680 (1994)).

5 The polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. In addition, as described in detail below, the polypeptides of the present invention can be used to raise polyclonal and monoclonal antibodies which are useful in a variety of assays for detecting 10 protein expression, localization, detection of interactions with other molecules, or for the isolation of a polypeptide (including a fusion polypeptide) of the invention.

15 In another aspect, the present invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention, which may be used to raise antibodies, particularly monoclonal antibodies, that bind specifically to a one or more of the polypeptides of the invention. The epitope of this polypeptide portion is an immunogenic or antigenic epitope of a polypeptide of the invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein is the immunogen. These immunogenic epitopes are believed to be confined to a few loci on the molecule. 20 On the other hand, a region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes (see, e.g., Geysen et al., *Proc. Natl. Acad. Sci. USA* 81:3998- 4002 (1983)).

25 As to the selection of peptides or polypeptides bearing an antigenic epitope (*i.e.*, that contain a region of a protein molecule to which an antibody can bind), it is well-known in the art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein (see, e.g., Sutcliffe, J.G., et al., *Science* 219:660-666 (1983)). Peptides capable of eliciting protein-reactive sera are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are not confined to the immunodominant regions of intact proteins (*i.e.*, immunogenic epitopes) or to the amino or carboxy

termini. Peptides that are extremely hydrophobic and those of six or fewer residues generally are ineffective at inducing antibodies that bind to the mimicked protein; longer peptides, especially those containing proline residues, usually are effective (Sutcliffe, J.G., *et al.*, *Science* 219:660-666 (1983)).

Epitope-bearing peptides and polypeptides of the invention designed according to the above guidelines preferably contain a sequence of at least five, more preferably at least seven or more amino acids contained within the amino acid sequence of a polypeptide of the invention. However, peptides or polypeptides comprising a larger portion of an amino acid sequence of a polypeptide of the invention, containing about 30 to about 50 amino acids, or any length up to and including the entire amino acid sequence of a given polypeptide of the invention, also are considered epitope-bearing peptides or polypeptides of the invention and also are useful for inducing antibodies that react with the mimicked protein. Preferably, the amino acid sequence of the epitope-bearing peptide is selected to provide substantial solubility in aqueous solvents (*i.e.*, the sequence includes relatively hydrophilic residues and highly hydrophobic sequences are preferably avoided); sequences containing proline residues are particularly preferred.

Non-limiting examples of epitope-bearing polypeptides or peptides that can be used to generate antibodies specific for the polypeptides of the invention include certain epitope-bearing regions of the polypeptides comprising amino acid sequences encoded by polynucleotides comprising one or more of the recombination site-encoding nucleic acid molecules of the invention, including those encoding *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* having the nucleotide sequences set forth in Figure 9 (or a nucleotide sequence complementary thereto); the complete amino acid sequences encoded by the three reading frames of the polynucleotides contained in the deposited clones described herein; and the amino acid sequences encoded by all reading frames of polynucleotides which hybridize under stringent hybridization conditions to polynucleotides having the nucleotide sequences encoding the recombination site sequences (or portions thereof) of the invention as set forth in Figure 9 (or a nucleic acid sequence complementary thereto). Other epitope-bearing polypeptides or peptides that may be used to generate antibodies specific for the polypeptides

of the invention will be apparent to one of ordinary skill in the art based on the primary amino acid sequences of the polypeptides of the invention described herein, via the construction of Kyte-Doolittle hydrophilicity and Jameson-Wolf antigenic index plots of the polypeptides of the invention using, for example, PROTEAN computer software (DNASTAR, Inc.; Madison, Wisconsin).

The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means for making peptides or polypeptides including recombinant means using nucleic acid molecules of the invention. For instance, a short epitope-bearing amino acid sequence may be fused to a larger polypeptide which acts as a carrier during recombinant production and purification, as well as during immunization to produce anti-peptide antibodies. Epitope-bearing peptides also may be synthesized using known methods of chemical synthesis (see, e.g., U.S. Patent No. 4,631,211 and Houghten, R. A., *Proc. Natl. Acad. Sci. USA* 82:5131-5135 (1985), both of which are incorporated by reference herein in their entireties).

As one of skill in the art will appreciate, the polypeptides of the present invention and epitope-bearing fragments thereof may be immobilized onto a solid support, by techniques that are well-known and routine in the art. By "solid support" is intended any solid support to which a peptide can be immobilized. Such solid supports include, but are not limited to nitrocellulose, diazocellulose, glass, polystyrene, polyvinylchloride, polypropylene, polyethylene, dextran, Sepharose, agar, starch, nylon, beads and microtitre plates. Linkage of the peptide of the invention to a solid support can be accomplished by attaching one or both ends of the peptide to the support. Attachment may also be made at one or more internal sites in the peptide. Multiple attachments (both internal and at the ends of the peptide) may also be used according to the invention. Attachment can be via an amino acid linkage group such as a primary amino group, a carboxyl group, or a sulphhydryl (SH) group or by chemical linkage groups such as with cyanogen bromide (CNBr) linkage through a spacer. For non-covalent attachments to the support, addition of an affinity tag sequence to the peptide can be used such as GST (Smith, D.B., and Johnson, K.S., *Gene* 67:31 (1988)), polyhistidines (Hochuli, E., *et al.*, *J. Chromatog.* 411:77 (1987)), or biotin. Such affinity tags

may be used for the reversible attachment of the peptide to the support. Such immobilized polypeptides or fragments may be useful, for example, in isolating antibodies directed against one or more of the polypeptides of the invention, or other proteins or peptides that recognize other proteins or peptides that bind to one or more of the polypeptides of the invention, as described below.

As one of skill in the art will also appreciate, the polypeptides of the present invention and the epitope-bearing fragments thereof described herein can be combined with one or more fusion partner proteins or peptides, or portions thereof, including but not limited to GST, His₆, Trx, and portions of the constant domain of immunoglobulins (Ig), resulting in chimeric or fusion polypeptides. These fusion polypeptides facilitate purification of the polypeptides of the invention (EP 0 394 827; Traunecker *et al.*, *Nature* 331:84- 86 (1988)) for use in analytical or diagnostic (including high-throughput) format.

15 *Antibodies*

In another aspect, the invention relates to antibodies that recognize and bind to the polypeptides (or epitope-bearing fragments thereof) or nucleic acid molecules (or portions thereof) of the invention. In a related aspect, the invention relates to antibodies that recognize and bind to one or more polypeptides encoded by all reading frames of one or more recombination site nucleic acid sequences or portions thereof, or to one or more nucleic acid molecules comprising one or more recombination site nucleic acid sequences or portions thereof, including but not limited to *att* sites (including *att*B1, *att*B2, *att*P1, *att*P2, *att*L1, *att*L2, *att*R1, *att*R2 and the like), *lox* sites (e.g., *lox*P, *lox*P511, and the like), FRT, and the like, or mutants, fragments, variants and derivatives thereof. See generally U.S. Patent No. 5,888,732, which is incorporated herein by reference in its entirety. The antibodies of the present invention may be polyclonal or monoclonal, and may be prepared by any of a variety of methods and in a variety of species according to methods that are well-known in the art. See, for instance, U.S. Patent No. 5,587,287; Sutcliffe, J.G., *et al.*, *Science* 219:660-666 (1983); Wilson *et al.*, *Cell* 37: 767 (1984); and Bittle, F.J., *et al.*, *J. Gen. Virol.* 66:2347-2354 (1985). Antibodies specific for any of the polypeptides or nucleic acid molecules described

herein, such as antibodies specifically binding to one or more of the polypeptides encoded by the recombination site nucleotide sequences, or one or more nucleic acid molecules, described herein or contained in the deposited clones, antibodies against fusion polypeptides (e.g., binding to fusion polypeptides between one or more of the fusion partner proteins and one or more of the recombination site polypeptides of the invention, as described herein), and the like, can be raised against the intact polypeptides or polynucleotides of the invention or one or more antigenic polypeptide fragments thereof.

As used herein, the term "antibody" (Ab) may be used interchangeably with the terms "polyclonal antibody" or "monoclonal antibody" (mAb), except in specific contexts as described below. These terms, as used herein, are meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to a polypeptide or nucleic acid molecule of the invention or a portion thereof. It will therefore be appreciated that, in addition to the intact antibodies of the invention, Fab, F(ab')₂ and other fragments of the antibodies described herein, and other peptides and peptide fragments that bind one or more polypeptides or polynucleotides of the invention, are also encompassed within the scope of the invention. Such antibody fragments are typically produced by proteolytic cleavage of intact antibodies, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Antibody fragments, and peptides or peptide fragments, may also be produced through the application of recombinant DNA technology or through synthetic chemistry.

Epitope-bearing peptides and polypeptides, and nucleic acid molecules or portions thereof, of the invention may be used to induce antibodies according to methods well known in the art, as generally described herein (see, e.g., Sutcliffe, *et al.*, *supra*; Wilson, *et al.*, *supra*; and Bittle, F. J., *et al.*, *J. Gen. Virol.* 66:2347-2354 (1985)).

Polyclonal antibodies according to this aspect of the invention may be made by immunizing an animal with one or more of the polypeptides or nucleic acid molecules of the invention described herein or portions thereof according to standard techniques (see, e.g., Harlow, E., and Lane, D., *Antibodies: A*

Laboratory Manual, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (1988); Kaufman, P.B., et al., In: *Handbook of Molecular and Cellular Methods in Biology and Medicine*, Boca Raton, Florida: CRC Press, pp. 468-469 (1995)). For producing antibodies that recognize and bind to the polypeptides or nucleic acid molecules of the invention or portions thereof, animals may be immunized with free peptide or free nucleic acid molecules; however, antibody titer may be boosted by coupling of the peptide to a macromolecular carrier, such as albumin, KLH, or tetanus toxoid (particularly for producing antibodies against the nucleic acid molecules of the invention or portions thereof; see Harlow and Lane, *supra*, at page 154), or to a solid phase carrier such as a latex or glass microbead. For instance, peptides containing cysteine may be coupled to carrier using a linker such as m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carrier using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice may be immunized with either free (if the polypeptide immunogen is larger than about 25 amino acids in length) or carrier-coupled peptides or nucleic acid molecules, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 µg peptide, polynucleotide, or carrier protein, and Freund's adjuvant. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of antibody which can be detected, for example, by ELISA assay using free peptide or nucleic acid molecule adsorbed to a solid surface. In another approach, cells expressing one or more of the polypeptides or polynucleotides of the invention or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies, according to routine immunological methods. In yet another method, a preparation of one or more of the polypeptides or polynucleotides of the invention is prepared and purified as described herein, to render it substantially free of natural contaminants. Such a preparation may then be introduced into an animal in order to produce polyclonal antisera of greater specific activity. The titer of antibodies in serum from an immunized animal, regardless of the method of immunization used, may be increased by selection of anti-peptide or anti-polynucleotide antibodies, for

instance, by adsorption to the peptide or polynucleotide on a solid support and elution of the selected antibodies according to methods well known in the art.

In an alternative method, the antibodies of the present invention are monoclonal antibodies (or fragments thereof which bind to one or more of the polypeptides of the invention). Such monoclonal antibodies can be prepared using hybridoma technology (Kohler *et al.*, *Nature* 256:495 (1975); Köhler *et al.*, *Eur. J. Immunol.* 6:511 (1976); Köhler *et al.*, *Eur. J. Immunol.* 6:292 (1976); Hammerling *et al.*, In: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., pp. 563-681 (1981)). In general, such procedures involve immunizing an animal (preferably a mouse) with a polypeptide or polynucleotide of the invention (or a fragment thereof), or with a cell expressing a polypeptide or polynucleotide of the invention (or a fragment thereof). The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP₂O), available from the American Type Culture Collection, Rockville, Maryland. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands *et al.* (*Gastroenterol.* 80:225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding one or more of the polypeptides or nucleic acid molecules of the invention, or fragments thereof. Hence, the present invention also provides hybridoma cells and cell lines producing monoclonal antibodies of the invention, particularly that recognize and bind to one or more of the polypeptides or nucleic acid molecules of the invention.

Alternatively, additional antibodies capable of binding to one or more of the polypeptides of the invention, or fragments thereof, may be produced in a two-step procedure through the use of anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and that, therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, antibodies specific for one or more of the polypeptides or polynucleotides of the invention, prepared as described above, are used to immunize an animal, preferably a mouse. The splenocytes of such an

-90-

animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to an antibody specific for one or more of the polypeptides or polynucleotides of the invention can be blocked by polypeptides of the invention themselves. Such antibodies comprise anti-idiotypic antibodies to the antibodies recognizing one or more of the polypeptides or polynucleotides of the invention, and can be used to immunize an animal to induce formation of further antibodies specific for one or more of the polypeptides or polynucleotides of the invention.

For use, the antibodies of the invention may optionally be detectably labeled by covalent or non-covalent attachment of one or more labels, including but not limited to chromogenic, enzymatic, radioisotopic, isotopic, fluorescent, toxic, chemiluminescent, or nuclear magnetic resonance contrast agents or other labels.

Examples of suitable enzyme labels include malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast-alcohol dehydrogenase, alpha-glycerol phosphate dehydrogenase, triose phosphate isomerase, peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase, and acetylcholine esterase.

Examples of suitable radioisotopic labels include ^3H , ^{111}In , ^{125}I , ^{131}I , ^{32}P , ^{35}S , ^{14}C , ^{51}Cr , ^{57}Co , ^{59}Fe , ^{75}Se , ^{152}Eu , ^{90}Y , ^{67}Cu , ^{217}Cl , ^{211}At , ^{212}Pb , ^{47}Sc , ^{109}Pd , etc.

^{111}In is a preferred isotope where in vivo imaging is used since it avoids the problem of dehalogenation of the ^{125}I or ^{131}I -labeled monoclonal antibody by the liver. In addition, this radionucleotide has a more favorable gamma emission energy for imaging (Perkins *et al.*, *Eur. J. Nucl. Med.* 10:296-301 (1985); Carasquillo *et al.*, *J. Nucl. Med.* 28:281-287 (1987)). For example, ^{111}In coupled to monoclonal antibodies with 1-(P-isothiocyanatobenzyl)-DPTA has shown little uptake in non-tumorous tissues, particularly the liver, and therefore enhances specificity of tumor localization (Esteban *et al.*, *J. Nucl. Med.* 28:861-870 (1987)).

Examples of suitable non-radioactive isotopic labels include ^{157}Gd , ^{55}Mn , ^{162}Dy , ^{52}Tr , and ^{56}Fe .

Examples of suitable fluorescent labels include an ^{152}Eu label, a fluorescein label, an isothiocyanate label, a rhodamine label, a phycoerythrin label, a

phycocyanin label, an allophycocyanin label, an o-phthaldehyde label, a green fluorescent protein (GFP) label, and a fluorescamine label.

Examples of suitable toxin labels include diphtheria toxin, ricin, and cholera toxin.

5 Examples of chemiluminescent labels include a luminal label, an isoluminal label, an aromatic acridinium ester label, an imidazole label, an acridinium salt label, an oxalate ester label, a luciferin label, a luciferase label, and an aequorin label.

10 Examples of nuclear magnetic resonance contrasting agents include heavy metal nuclei such as Gd, Mn, and iron.

15 Typical techniques for binding the above-described labels to the antibodies of the invention are provided by Kennedy *et al.*, *Clin. Chim. Acta* 70:1-31 (1976), and Schurs *et al.*, *Clin. Chim. Acta* 81:1-40 (1977). Coupling techniques mentioned in the latter are the glutaraldehyde method, the periodate method, the dimaleimide method, the m-maleimidobenzyl-N-hydroxy-succinimide ester method, all of which methods are incorporated by reference herein.

20 It will be appreciated by one of ordinary skill that the antibodies of the present invention may alternatively be coupled to a solid support, to facilitate, for example, chromatographic and other immunological procedures using such solid phase-immobilized antibodies. Included among such procedures are the use of the antibodies of the invention to isolate or purify polypeptides comprising one or more epitopes encoded by the nucleic acid molecules of the invention (which may be fusion polypeptides or other polypeptides of the invention described herein), or to isolate or purify polynucleotides comprising one or more recombination site sequences of the invention or portions thereof. Methods for isolation and purification of polypeptides (and, by analogy, polynucleotides) by affinity chromatography, for example using the antibodies of the invention coupled to a solid phase support, are well-known in the art and will be familiar to one of ordinary skill. The antibodies of the invention may also be used in other applications, for example to cross-link or couple two or more proteins, polypeptides, polynucleotides, or portions thereof into a structural and/or functional complex. In one such use, an antibody of the invention may have two

or more distinct epitope-binding regions that may bind, for example, a first polypeptide (which may be a polypeptide of the invention) at one epitope-binding region on the antibody and a second polypeptide (which may be a polypeptide of the invention) at a second epitope-binding region on the antibody, thereby bringing the first and second polypeptides into close proximity to each other such that the first and second polypeptides are able to interact structurally and/or functionally (as, for example, linking an enzyme and its substrate to carry out enzymatic catalysis, or linking an effector molecule and its receptor to carry out or induce a specific binding of the effector molecule to the receptor or a response to the effector molecule mediated by the receptor). Additional applications for the antibodies of the invention include, for example, the preparation of large-scale arrays of the antibodies, polypeptides, or nucleic acid molecules of the invention, or portions thereof, on a solid support, for example to facilitate high-throughput screening of protein or RNA expression by host cells containing nucleic acid molecules of the invention (known in the art as "chip array" protocols; see, e.g., U.S. Patent Nos. 5,856,101, 5,837,832, 5,770,456, 5,744,305, 5,631,734, and 5,593,839, which are directed to production and use of chip arrays of polypeptides (including antibodies) and polynucleotides, and the disclosures of which are incorporated herein by reference in their entireties). By "solid support" is intended any solid support to which an antibody can be immobilized. Such solid supports include, but are not limited to nitrocellulose, diazocellulose, glass, polystyrene, polyvinylchloride, polycarbonate, polypropylene, polyethylene, dextran, Sepharose, agar, starch, nylon, beads and microtitre plates. Preferred are beads made of glass, latex or a magnetic material. Linkage of an antibody of the invention to a solid support can be accomplished by attaching one or both ends of the antibody to the support. Attachment may also be made at one or more internal sites in the antibody. Multiple attachments (both internal and at the ends of the antibody) may also be used according to the invention. Attachment can be via an amino acid linkage group such as a primary amino group, a carboxyl group, or a sulphhydryl (SH) group or by chemical linkage groups such as with cyanogen bromide (CNBr) linkage through a spacer. For non-covalent attachments, addition of an affinity tag sequence to the peptide can be used such as GST

-93-

(Smith, D.B., and Johnson, K.S., *Gene* 67:31 (1988)), polyhistidines (Hochuli, E., et al., *J. Chromatog.* 411:77 (1987)), or biotin. Alternatively, attachment can be accomplished using a ligand which binds the Fc region of the antibodies of the invention, e.g., protein A or protein G. Such affinity tags may be used for the reversible attachment of the antibodies to the support. Peptides may also be recognized via specific ligand-receptor interactions or using phage display methodologies that will be familiar to the skilled artisan, for their ability to bind polypeptides of the invention or fragments thereof.

5

10

Kits

In another aspect, the invention provides kits which may be used in producing the nucleic acid molecules, polypeptides, vectors, host cells, and antibodies, and in the recombinational cloning methods, of the invention. Kits according to this aspect of the invention may comprise one or more containers, which may contain one or more of the nucleic acid molecules, primers, polypeptides, vectors, host cells, or antibodies of the invention. In particular, a kit of the invention may comprise one or more components (or combinations thereof) selected from the group consisting of one or more recombination proteins (e.g., Int) or auxiliary factors (e.g. IHF and/or Xis) or combinations thereof, one or more compositions comprising one or more recombination proteins or auxiliary factors or combinations thereof (for example, GATEWAY™ LR Clonase™ Enzyme Mix or GATEWAY™ BP Clonase™ Enzyme Mix) one or more Destination Vector molecules (including those described herein), one or more Entry Clone or Entry Vector molecules (including those described herein), one or more primer nucleic acid molecules (particularly those described herein), one or more host cells (e.g. competent cells, such as *E. coli* cells, yeast cells, animal cells (including mammalian cells, insect cells, nematode cells, avian cells, fish cells, etc.), plant cells, and most particularly *E. coli* DB3, DB3.1 (preferably *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells; Life Technologies, Inc., Rockville, MD), DB4 and DB5; see U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, and the corresponding U.S. Utility Application No. _____ of Hartley et al., entitled "Cells Resistant to Toxic Genes and Uses Thereof," filed

15

20

25

30

on even day herewith, the disclosures of which are incorporated by reference herein in its entirety), and the like. In related aspects, the kits of the invention may comprise one or more nucleic acid molecules encoding one or more recombination sites or portions thereof, such as one or more nucleic acid molecules comprising a nucleotide sequence encoding the one or more recombination sites (or portions thereof) of the invention, and particularly one or more of the nucleic acid molecules contained in the deposited clones described herein. Kits according to this aspect of the invention may also comprise one or more isolated nucleic acid molecules of the invention, one or more vectors of the invention, one or more primer nucleic acid molecules of the invention, and/or one or more antibodies of the invention. The kits of the invention may further comprise one or more additional containers containing one or more additional components useful in combination with the nucleic acid molecules, polypeptides, vectors, host cells, or antibodies of the invention, such as one or more buffers, one or more detergents, one or more polypeptides having nucleic acid polymerase activity, one or more polypeptides having reverse transcriptase activity, one or more transfection reagents, one or more nucleotides, and the like. Such kits may be used in any process advantageously using the nucleic acid molecules, primers, vectors, host cells, polypeptides, antibodies and other compositions of the invention, for example in methods of synthesizing nucleic acid molecules (e.g., via amplification such as via PCR), in methods of cloning nucleic acid molecules (preferably via recombinational cloning as described herein), and the like.

Optimization of Recombinational Cloning System

The usefulness of a particular nucleic acid molecule, or vector comprising a nucleic acid molecule, of the invention in methods of recombinational cloning may be determined by any one of a number of assay methods. For example, Entry and Destination vectors of the present invention may be assessed for their ability to function (*i.e.*, to mediate the transfer of a nucleic acid molecule, DNA segment, gene, cDNA molecule or library from a cloning vector to an Expression Vector) by carrying out a recombinational cloning reaction as described in more detail in the Examples below and as described in U.S. Application Nos. 08/663,002, filed

-95-

June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, 09/177,387, filed October 23, 1998, and 60/108,324, filed November 13, 1998, the disclosures of which are incorporated by reference herein in their entireties. Alternatively, the functionality of Entry and Destination Vectors prepared according to the invention may be assessed by examining the ability of these vectors to recombine and create cointegrate molecules, or to transfer a nucleic acid molecule of interest, using an assay such as that described in detail below in Example 19. Analogously, the formulation of compositions comprising one or more recombination proteins or combinations thereof, for example GATEWAY™ LR Clonase™ Enzyme Mix and GATEWAY™ BP Clonase™ Enzyme Mix, may be optimized using assays such as those described below in Example 18.

Uses

There are a number of applications for the compositions, methods and kits of the present invention. These uses include, but are not limited to, changing vectors, targeting gene products to intracellular locations, cleaving fusion tags from desired proteins, operably linking nucleic acid molecules of interest to regulatory genetic sequences (e.g., promoters, enhancers, and the like), constructing genes for fusion proteins, changing copy number, changing replicons, cloning into phages, and cloning, e.g., PCR products, genomic DNAs, and cDNAs. In addition, the nucleic acid molecules, vectors, and host cells of the invention may be used in the production of polypeptides encoded by the nucleic acid molecules, in the production of antibodies directed against such polypeptides, in recombinational cloning of desired nucleic acid sequences, and in other applications that may be enhanced or facilitated by the use of the nucleic acid molecules, vectors, and host cells of the invention.

In particular, the nucleic acid molecules, vectors, host cells, polypeptides, antibodies, and kits of the invention may be used in methods of transferring one or more desired nucleic acid molecules or DNA segments, for example one or more genes, cDNA molecules or cDNA libraries, into a cloning or Expression Vector for use in transforming additional host cells for use in cloning or

amplification of, or expression of the polypeptide encoded by, the desired nucleic acid molecule or DNA segment. Such recombinational cloning methods which may advantageously use the nucleic acid molecules, vectors, and host cells of the invention, are described in detail in the Examples below, and in commonly owned U.S. Application Nos. 08/486,139, filed June 7, 1995, 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, 09/177,387, filed October 23, 1998, and 60/108,324, filed November 13, 1998, the disclosures of all of which are incorporated by reference herein in their entireties.

5

10

It will be understood by one of ordinary skill in the relevant arts that other suitable modifications and adaptations to the methods and applications described herein are readily apparent from the description of the invention contained herein in view of information known to the ordinarily skilled artisan, and may be made without departing from the scope of the invention or any embodiment thereof. Having now described the present invention in detail, the same will be more clearly understood by reference to the following examples, which are included herewith for purposes of illustration only and are not intended to be limiting of the invention.

15

20

Examples

Example 1: Recombination Reactions of Bacteriophage λ

The *E. coli* bacteriophage λ can grow as a lytic phage, in which case the host cell is lysed, with the release of progeny virus. Alternatively, lambda can integrate into the genome of its host by a process called lysogenization (see Figure 60). In this lysogenic state, the phage genome can be transmitted to daughter cells for many generations, until conditions arise that trigger its excision from the genome. At this point, the virus enters the lytic part of its life cycle. The control of the switch between the lytic and lysogenic pathways is one of the best understood processes in molecular biology (M. Ptashne, *A Genetic Switch*, Cell Press, 1992).

30

The integrative and excisive recombination reactions of λ , performed *in vitro*, are the basis of Recombinational Cloning System of the present invention. They can be represented schematically as follows:

5



10

The four att sites contain binding sites for the proteins that mediate the reactions. The wild type attP, attB, attL, and attR sites contain about 243, 25, 100, and 168 base pairs, respectively. The attB x attP reaction (hereinafter referred to as a "BP Reaction," or alternatively and equivalently as an "Entry Reaction" or a "Gateward Reaction") is mediated by the proteins Int and IHF. The attL x attR reaction (hereinafter referred to as an "LR Reaction," or alternatively and equivalently as a "Destination Reaction") is mediated by the proteins Int, IHF, and Xis. Int (integrase) and Xis (excisionase) are encoded by the λ genome, while IHF (integration host factor) is an *E. coli* protein. For a general review of lambda recombination, see: A. Landy, *Ann. Rev. Biochem.* 58: 913-949 (1989).

15

20

Example 2: Recombination Reactions of the Recombinational Cloning System

The LR Reaction -- the exchange of a DNA segment from an Entry Clone to a Destination Vector -- is the *in vitro* version of the λ excision reaction:

25



30

There is a practical imperative for this configuration: after an LR Reaction in one configuration of the present method, an att site usually separates a functional motif (such as a promoter or a fusion tag) from a nucleic acid molecule of interest in an Expression Clone, and the 25 bp attB site is much smaller than the attP, attL, and attR sites.

Note that the recombination reaction is conservative, i.e., there is no net synthesis or loss of base pairs. The DNA segments that flank the recombination

sites are merely switched. The wild type λ recombination sites are modified for purposes of the GATEWAYTM Cloning System, as follows:

To create certain preferred Destination Vectors, a part (43 bp) of attR was removed, to make the excisive reaction irreversible and more efficient (W. Bushman et al., *Science* 230: 906, 1985). The attR sites in preferred Destination Vectors of the invention are 125 bp in length. Mutations were made to the core regions of the att sites, for two reasons: (1) to eliminate stop codons, and (2) to ensure specificity of the recombination reactions (i.e., attR1 reacts only with attL1, attR2 reacts only with attL2, etc.).

Other mutations were introduced into the short (5 bp) regions flanking the 15 bp core regions of the attB sites to minimize secondary structure formation in single-stranded forms of attB plasmids, e.g., in phagemid ssDNA or in mRNA. Sequences of attB1 and attB2 to the left and right of a nucleic acid molecule of interest after it has been cloned into a Destination Vector are given in Figure 6.

Figure 61 illustrates how an Entry Clone and a Destination Vector recombine in the LR Reaction to form a co-integrate, which resolves through a second reaction into two daughter molecules. The two daughter molecules have the same general structure regardless of which pair of sites, attL1 and attR1 or attL2 and attR2, react first to form the co-integrate. The segments change partners by these reactions, regardless of whether the parental molecules are both circular, one is circular and one is linear, or both are linear. In this example, selection for ampicillin resistance carried on the Destination Vector, which also carries the death gene ccdB, provides the means for selecting only for the desired attB product plasmid.

Example 3: Protein Expression in the Recombinational Cloning System

Proteins are expressed *in vivo* as a result of two processes, transcription (DNA into RNA), and translation (RNA into protein). For a review of protein expression in prokaryotes and eukaryotes, see Example 13 below. Many vectors (pUC, BlueScript, pGem) use interruption of a transcribed lacZ gene for blue-white screening. These plasmids, and many Expression Vectors, use the lac promoter to control expression of cloned genes. Transcription from the lac

promoter is turned on by adding the inducer IPTG. However, a low level of RNA is made in the absence of inducer, i.e., the lac promoter is never completely off. The result of this "leakiness" is that genes whose expression is harmful to *E. coli* may prove difficult or impossible to clone in vectors that contain the lac promoter, or they may be cloned only as inactive mutants.

In contrast to other gene expression systems, nucleic acid molecules cloned into an Entry Vector may be designed *not* to be expressed. The presence of the strong transcriptional terminator *rrnB* (Orosz, et al., *Eur. J. Biochem.* 201: 653, 1991) just upstream of the attL1 site keeps transcription from the vector promoters (drug resistance and replication origin) from reaching the cloned gene. However, if a toxic gene is cloned into a Destination Vector, the host may be sick, just as in other expression systems. But the reliability of subcloning by *in vitro* recombination makes it easier to recognize that this has happened -- and easier to try another expression option in accordance with the methods of the invention, if necessary.

Example 4: Choosing the Right Entry Vector

There are two kinds of choices that must be made in choosing the best Entry Vector, dictated by (1) the particular DNA segment that is to be cloned, and (2) what is to be accomplished with the cloned DNA segment. These factors are critical in the choice of Entry Vector used, because when the desired nucleic acid molecule of interest is moved from the Entry Vector to a Destination Vector, all the base pairs between the nucleic acid molecule of interest and the Int cutting sites in attL1 and attL2 (such as in Figure 6) move into the Destination Vector as well. For genomic DNAs that are not expressed as a result of moving into a Destination Vector, these decisions are not as critical.

For example, if an Entry Vector with certain translation start signals is used, those sequences will be translated into amino acids if an amino-terminal fusion to the desired nucleic acid molecule of interest is made. Whether the desired nucleic acid molecule of interest is to be expressed as fusion protein, native protein, or both, dictates whether translational start sequences must be included between the attB sites of the clone (native protein) or, alternatively, supplied by the Destination

-100-

Vector (fusion protein). In particular, Entry Clones that include translational start sequences may prove less suitable for making fusion proteins, as internal initiation of translation at these sites can decrease the yield of N-terminal fusion protein. These two types of expression afforded by the compositions and methods of the invention are illustrated in Figure 62.

No Entry Vector is likely to be optimal for all applications. The nucleic acid molecule of interest may be cloned into any of several optimal Entry Vectors.

As an example, consider pENTR7 (Figure 16) and pENTR11 (Figure 20), which are useful in a variety of applications, including (but not limited to):

10

- Cloning cDNAs from most of the commercially available libraries. The sites to the left and right of the *ccdB* death gene have been chosen so that directional cloning is possible if the DNA to be cloned does not have two or more of these restriction sites.

15

- Cloning of genes directionally: *SalI*, *BamHI*, *XmnI* (blunt), or *KpnI* on the left of *ccdB*; *NotI*, *XhoI*, *XbaI*, or *EcoRV* (blunt), on the right.

20

- Cloning of genes or gene fragments with a blunt amino end at the *XmnI* site. The *XmnI* site has four of the six most favored bases for eukaryotic expression (see Example 13, below), so that if the first three bases of the DNA to be cloned are ATG, the open reading frame (ORF) will be expressed in eukaryotic cells (e.g., mammalian cells, insect cells, yeast cells) when it is transcribed in the appropriate Destination Vector. In addition, in pENTR11, a Shine-Dalgarno sequence is situated 8 bp upstream, for initiating protein synthesis in a prokaryotic host cell (particularly a bacterial cell, such as *E. coli*) at an ATG.

25

- Cleaving off amino terminal fusions (e.g., His₆, GST, or thioredoxin) using the highly specific TEV (Tobacco Etch Virus) protease (available from Life Technologies, Inc.). If the nucleic acid molecule of interest is cloned at the

30

-101-

blunt *XmnI* site, TEV cleavage will leave two amino acids on the amino end of the expressed protein.

5 • Selecting against uncut or singly cut Entry Vector molecules during cloning with restriction enzymes and ligase. If the *ccdB* gene is not removed with a double digest, it will kill any recipient *E. coli* cell that does not contain a mutation that makes the cell resistant to *ccdB* (see U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, the disclosure of which is incorporated by reference herein in its entirety).

10

• Allowing production of amino fusions with ORFs in all cloning sites. There are no stop codons (in the *attL1* reading frame) upstream of the *ccdB* gene.

In addition, pENTR11 is also useful in the following applications:

15

• Cloning cDNAs that have an *NcoI* site at the initiating ATG into the *NcoI* site. Similar to the *XmnI* site, this site has four of the six most favored bases for eukaryotic expression. Also, a Shine-Dalgarno sequence is situated 8 bp upstream, for initiating protein synthesis in a prokaryotic host cell (particularly a bacterial cell, such as *E. coli*) at an ATG.

20

• Producing carboxy fusion proteins with ORFs positioned in phase with the reading frame convention for carboxy-terminal fusions (see Figure 20A).

25

Table 1 lists some non-limiting examples of Entry Vectors and their characteristics, and Figures 10-20 show their cloning sites. All of the Entry Vectors listed in Table 1 are available commercially from Life Technologies, Inc., Rockville, Maryland. Other Entry Vectors not specifically listed here, which comprise alternative or additional features may be made by one of ordinary skill using routine methods of molecular and cellular biology, in view of the disclosure contained herein.

30

Table 1 Examples of Entry Vectors

Designation	Mnemonic Name	Class of Entry Vector	Distinctive Cloning Sites	Amino Fusions	Native Protein in E. coli	Native Protein in Eukaryotic Cells	Protein Synthesis Features
pENTR-1A, 2B, 3C	Minimal blunt RF A, B, C	Alternative Reading Frame Vectors	Reading frame A, B, or C; blunt cut closest to attL1	Good	Poor	Good	Minimal amino acids between tag and protein; no SD
pENTR4	Minimal Nco	Restr. Enz. Cleavage Vectors	Nco I site (common in euk. cDNAs) closest to attL1	Good	Poor	Good	Good Kozac; no SD
pENTR5	Minimal Nde	Restr. Enz. Cleavage Vectors	NdeI site closest to attL1	Good	Poor	Poor at Nde I, Good at Xmn I	No SD; poor Kozac at Nde, good at Xmn
pENTR6	Minimal Sph	Restr. Enz. Cleavage Vectors	Sph I site closest to attL1	Good	Poor	Poor at Sph I, Good at Xmn I	No SD; poor Kozac at Sph, good at Xmn
pENTR7	TEV Blunt	TEV Cleavage Site Present	Xmn I (blunt) is first cloning site after TEV site	Good	Poor	Good at Xmn I site	TEV protease leaves Gly-Thr on amino end of protein; no SD
pENTR8	TEV Nco	TEV Cleavage Site Present	Nco I is first cloning site after TEV site	Good	Poor	Good	TEV protease leaves Gly-Thr on amino end of protein; no SD

pENTR9	TEV Nde	TEV Cleavage Site Present	Nde I is first cloning site after TEV site	Good	Poor	Poor	TEV protease leaves Gly-Thr on amino end of protein; no SD, poor Kozac
pENTR10	Nde with SD	Good SD for E.coli Expression	Strong SD; Nde I site, no TEV	Poor	Good	Poor	Strong SD, internal starts in amino fusions. Poor Kz. No TEV
pENTR11	2 X SD+Kozac	Good SD for E.coli Expression	Xmn I (blunt) and Nco I sites each preceded by SD and Kozac	Good	Good	Good	Strong SD/Koz Internal starts in amino fusions. No TEV

5

10

15

20

Entry vectors pENTR1A (Figures 10A and 10B), pENTR2B (Figures 11A and 11B), and pENTR3C (Figures 12A and 12B) are almost identical, except that the restriction sites are in different reading frames. Entry vectors pENTR4 (Figures 13A and 13B), pENTR5 (Figures 14A and 14B), and pENTR6 (Figures 15A and 15B) are essentially identical to pENTR1A, except that the blunt *Dra*I site has been replaced with sites containing the ATG methionine codon: *Nco*I in pENTR4, *Nde*I in pENTR5, and *Sph*I in pENTR6. Nucleic acid molecules that contain one of these sites at the initiating ATG can be conveniently cloned in these Entry vectors. The *Nco*I site in pENTR4 is especially useful for expression of nucleic acid molecules in eukaryotic cells, since it contains many of the bases that give efficient translation (see Example 13, below). (Nucleic acid molecules of interest cloned into the *Nde*I site of pENTR5 are not expected to be highly expressed in eukaryotic cells, because the cytosine at position -3 from the initiating ATG is rare in eukaryotic genes.)

Entry vectors pENTR7 (Figures 16A and 16B), pENTR8 (Figures 17A and 17B), and pENTR9 (Figures 18A and 18B) contain the recognition site for the TEV protease between the attL1 site and the cloning sites. Cleavage sites for *Xmn*I (blunt), *Nco*I, and *Nde*I, respectively, are the most 5' sites in these Entry vectors. Amino fusions can be removed efficiently if nucleic acid molecules are cloned into these Entry vectors. TEV protease is highly active and highly specific.

Example 5: Controlling Reading Frame

25

30

One of the trickiest tasks in expression of cloned nucleic acid molecules is making sure the reading frame is correct. (Reading frame is important if fusions are being made between two ORFs, for example between a nucleic acid molecule of interest and a His6 or GST domain.) For purposes of the present invention, the following convention has been adopted: The reading frame of the DNA cloned into any Entry Vector must be in phase with that of the attB1 site shown in Figure 16A, pENTR7. Notice that the six As of the attL1 site are split into two lysine codons (aaa aaa). The Destination Vectors that make amino fusions were constructed such that they enter the attR1 site in this reading frame.

-105-

Destination Vectors for carboxy terminal fusions were also constructed, including those containing His₆ (pDEST23; Figure 43), GST (pDEST24; Figure 44), or thioredoxin (pDEST25; Figure 45) C-terminal fusion sequences.

Therefore, if a nucleic acid molecule of interest is cloned into an Entry Vector so that the aaa aaa reading frame within the attL1 site is in phase with the nucleic acid molecule's ORF, amino terminal fusions will automatically be correctly phased, for all the fusion tags. This is a significant improvement over the usual case, where each different vector can have different restriction sites and different reading frames.

See Example 15 for a practical example of how to choose the most appropriate combinations of Entry Vector and Destination Vector.

Materials

Unless otherwise indicated, the following materials were used in the remaining Examples included herein:

5X LR Reaction Buffer:

200-250 mM (preferably 250 mM) Tris-HCl, pH 7.5

250-350 mM (preferably 320 mM) NaCl

1.25-5 mM (preferably 4.75 mM) EDTA

12.5-35 mM (preferably 22-35 mM, and most preferably 35 mM)

Spermidine-HCl

1 mg/ml bovine serum albumin

GATEWAY™ LR Clonase™ Enzyme Mix:

per 4 µl of 1X LR Reaction Buffer:

150 ng carboxy-His6-tagged Int (see U.S. Appl. Nos. 60/108,324, filed November 13, 1998, and 09/438,358, filed November 12, 1999, both entirely incorporated by reference herein)

-106-

25 ng carboxy-His6-tagged Xis (see U.S. Appl. Nos. 60/108,324, filed November 13, 1998, and 09/438,358, filed November 12, 1999, both entirely incorporated by reference herein)

30 ng IHF

5 50% glycerol

5X BP Reaction Buffer:

125 mM Tris-HCl, pH 7.5.

110 mM NaCl

10 25 mM EDTA

25 mM Spermidine-HCl

5 mg/ml bovine serum albumin

GATEWAY™ BP Clonase™ Enzyme Mix:

15 per 4 µl of 1X BP Reaction Buffer:

200 ng carboxy-His6-tagged Int (see U.S. Appl. Nos. 60/108,324, filed November 13, 1998, and 09/438,358, filed November 12, 1999, both entirely incorporated by reference herein)

80 ng IHF

20 50% glycerol

10X Clonase Stop Solution:

25 50 mM Tris-HCl, pH 8.0

1 mM EDTA

2 mg/ml Proteinase K

Example 6: LR ("Destination") Reaction

To create a new Expression Clone containing the nucleic acid molecule of interest (and which may be introduced into a host cell, ultimately for production of the polypeptide encoded by the nucleic acid molecule), an Entry Clone or Vector containing the nucleic acid molecule of interest, prepared as described

-107-

herein, is reacted with a Destination Vector. In the present example, a β -Gal gene flanked by attL sites is transferred from an Entry Clone to a Destination Vector.

Materials needed:

- 5 • 5 X LR Reaction buffer
- Destination Vector (preferably linearized), 75-150 ng/ μ l
- Entry Clone containing nucleic acid molecule of interest, 100-300 ng in \leq 8 μ l
TE buffer
- Positive control Entry Clone (pENTR- β -Gal) DNA (See note, below)
- 10 • Positive control Destination Vector, pDEST1 (pTrc), 75 ng/ μ l
- GATEWAY™ LR Clonase™ Enzyme Mix (stored at - 80° C)
- 10X Clonase Stop solution
- pUC19 DNA, 10 pg/ μ l
- Chemically competent *E. coli* cells (competence: $\geq 1 \times 10^7$ CFU/ μ g), 400 μ l.
- 15 • LB Plates containing ampicillin (100 μ g/ml) and methicillin (200 μ g/ml) \pm
X-gal and IPTG (See below)

Notes:

Preparation of the Entry Clone DNA: Miniprep DNA that has been treated with RNase works well. A reasonably accurate quantitation ($\pm 50\%$) of the DNA to be cloned is advised, as the GATEWAY™ reaction appears to have an optimum of about 100-300 ng of Entry Clone per 20 μ l of reaction mix.

The positive control Entry Clone, pENTR- β -Gal, permits functional analysis of clones based on the numbers of expected blue vs. white colonies on LB plates containing IPTG + Bluo-gal (or X-gal), in addition to ampicillin (100 μ g/ml) and methicillin (200 μ g/ml). Because β -Galactosidase is a large protein, it often yields a less prominent band than many smaller proteins do on SDS protein gels.

In the Positive Control Entry Vector pENTR- β -Gal, the coding sequence of β -Gal has been cloned into pENTR11 (Figures 20A and 20B), with translational start signals permitting expression in *E. coli*, as well as in eukaryotic

-108-

cells. The positive control Destination Vector, for example pDEST1 (Figure 21), is preferably linearized.

To prepare X-gal + IPTG plates, either of the following protocols may be used:

5

A. With a glass rod, spread over the surface of an LB agar plate: 40 µl of 20 mg/ml X-gal (or Bluo-gal) in DMF plus 4 µl 200 mg/ml IPTG. Allow liquid to adsorb into agar for 3-4 hours at 37° C before plating cells.

10

B. To liquid LB agar at ~45°C, add: X-gal (or Bluo-Gal) (20 mg/ml in DMF) to make 50 µg/ml and IPTG (200 mM in water) to make 0.5-1 mM, just prior to pouring plates. Store X-gal and Bluo-Gal in a light-shielded container.

15

Colony color may be enhanced by placing the plates at 5°C for a few hours after the overnight incubation at 37°C. Protocol B can give more consistent colony color than A, but A is more convenient when selection plates are needed on short notice.

20

Recombination in Clonase reactions continues for many hours. While incubations of 45-60 minutes are usually sufficient, reactions with large DNAs, or in which both parental DNAs are supercoiled, or which will be transformed into cells of low competence, can be improved with longer incubation times, such as 2-24 hours at 25°C.

Procedure:

25

1. Assemble reactions as follows (combine all components at room temperature, except GATEWAY™ LR Clonase™ Enzyme Mix ("Clonase LR"), before removing Clonase LR from frozen storage):

	Tube 1 Neg.	Tube 2 Pos.	Tube 3 Neg.	Tube 4 Test
5	p-Gate- β Gal, (Positive control Entry Clone) 75 ng/ μ l	4 μ l	4 μ l	
10	pDEST1 (Positive control Destination Vector), 75 ng/ μ l	4 μ l	4 μ l	
15	Your Entry Clone (100-300 ng)			1 - 8 μ l
20	Destination Vector for your nucleic acid molecule, 75 ng/ μ l			4 μ l
	5 X LR Reaction Buffer	4 μ l	4 μ l	4 μ l
	TE	8 μ l	4 μ l	To 20 μ l
	GATEWAY™ LR Clonase™ Enzyme Mix (store at - 80° C, add last)	---	4 μ l	---
	Total Volume	20 μ l	20 μ l	20 μ l

2. Remove the GATEWAY™ LR Clonase™ Enzyme Mix from the -80° C freezer, place immediately on ice. The Clonase takes only a few minutes to thaw.
3. Add 4 μ l of GATEWAY™ LR Clonase™ Enzyme Mix to reactions #2 and #4;
4. Return GATEWAY™ LR Clonase™ Enzyme Mix to - 80° C freezer.
5. Incubate tubes at 25° for at least 60 minutes.
6. Add 2 μ l Clonase Stop solution to all reactions. Incubate for 20 min at 37°C. (This step usually increases the total number of colonies obtained by 10-20 fold.)
7. Transform 2 μ l into 100 μ l competent *E. coli*. Select on plates containing ampicillin at 100 μ g/ml.

Example 7: Transformation of *E. coli*

To introduce cloning or Expression Vectors prepared using the recombinational cloning system of the invention, any standard *E. coli* transformation protocol should be satisfactory. The following steps are recommended for best results:

-110-

1. Let the mixture of competent cells and Recombinational Cloning System reaction product stand on ice at least 15 minutes prior to the heat-shock step. This gives time for the recombination proteins to dissociate from the DNA, and improves the transformation efficiency.

5

2. Expect the reaction to be about 1%-5% efficient, i.e., 2 µl of the reaction should contain at least 100 pg of the Expression Clone plasmid (taking into account the amounts of each parental plasmid in the reaction, and the subsequent dilution). If the E. coli cells have a competence of 10^7 CFU/µg, 100 pg of the desired clone plasmid will give about 1000 colonies, or more, if the entire transformation is spread on one ampicillin plate.
3. Always do a control pUC DNA transformation. If the number of colonies is not what you expect, the pUC DNA transformation gives you an indication of where the problem was.

10

15

Example 8: Preparation of attB-PCR Product

For preparation of attB-PCR products in the PCR cloning methods described in Example 9 below, PCR primers containing attB1 and attB2 sequences are used. The attB1 and attB2 primer sequences are as follows:

attB1: 5'-GGGGACAAGTTGTACAAAAAAGCAGGCT- (template-specific sequence)-3'

attB2: 5'-GGGGACCCTTGTACAAGAAAGCTGGGT- (template-specific sequence)-3'

The attB1 sequence should be added to the amino primer, and the attB2 sequence to the carboxy primer. The 4 guanines at the 5' ends of each of these primers enhance the efficiency of the minimal 25 bp attB sequences as substrates for use in the cloning methods of the invention.

Standard PCR conditions may be used to prepare the PCR product. The following suggested protocol employs PLATINUM *Taq* DNA Polymerase High

20

30

-111-

Fidelity®, available commercially from Life Technologies, Inc. (Rockville, MD). This enzyme mix eliminates the need for hot starts, has improved fidelity over Taq, and permits synthesis of a wide range of amplicon sizes, from 200 bp to 10 kb, or more, even on genomic templates.

5

Materials needed:

- PLATINUM Taq DNA Polymerase High Fidelity® (Life Technologies, Inc.)
- attB1- and attB2- containing primer pair (see above) specific for your template
- DNA template (linearized plasmid or genomic DNA)
- 10X High Fidelity PCR Buffer
- 10 mM dNTP mix
- PEG/MgCl₂ Mix (30% PEG 8000, 30 mM MgCl₂)

15

Procedure:

1.) Assemble the reaction as follows:

Component	Reaction with <u>Plasmid Target</u>	Reaction with <u>Genomic</u> Target
10X High Fidelity PCR Buffer	5 µl	5 µl
dNTP Mix 10 mM	1 µl	1 µl
MgSO ₄ , 50mM	2 µl	2 µl
attB1 Primer, 10 µM	2 µl	1 µl
attB2 Primer, 10 µM	2 µl	1 µl
Template DNA	1-5 ng*	≥ 100 ng
PLATINUM Taq High Fidelity	2 µl	1 µl
Water	to 50 µl	to 50 µl

* Use of higher amounts of plasmid template may permit fewer cycles (10-15) of PCR.

35

-112-

2.) Add 2 drops mineral oil, as appropriate.

3.) Denature for 30 sec. at 94°C.

4.) Perform 25 cycles:

5 94°C for 15 sec-30 sec

55°C for 15 sec-30 sec

68°C for 1 min per kb of template.

10 5.) Following the PCR reaction, apply 1-2 µl of the reaction mixture to an agarose gel, together with size standards (e.g., 1 Kb Plus Ladder, Life Technologies, Inc.) and quantitation standards (e.g., Low Mass Ladder, Life Technologies, Inc.), to assess the yield and uniformity of the product.

15 Purification of the PCR product is recommended, to remove attB primer dimers which can clone efficiently into the Entry Vector. The following protocol is fast and will remove DNA <300 bp in size:

16 6.) Dilute the 50 µl PCR reaction to 200 µl with TE.

17 7.) Add 100 µl PEG/MgCl₂ Solution. Mix and centrifuge immediately at 13,000 RPM for 10 min at room temperature. Remove the supernatant (pellet is clear and hard to see).

18 8.) Dissolve the pellet in 50 µl TE and check recovery on a gel.

25 If the starting PCR template is a plasmid that contains the gene for Kan^r, it is advisable to treat the completed PCR reaction with the restriction enzyme *Dpn*I, to degrade the plasmid since unreacted residual starting plasmid is a potential source of false-positive colonies from the transformation of the GATEWAY™ Cloning System reaction. Adding ~5 units of *Dpn*I to the completed PCR reaction and incubating for 15 min at 37°C will eliminate this potential problem. Heat inactivate the *Dpn*I at 65°C for 15 min, prior to using the PCR product in the GATEWAY™ Cloning System reaction.

Example 9: Cloning attB-PCR products into Entry Vectors via the BP ("Gateward") Reaction

The addition of 5'-terminal attB sequences to PCR primers allows synthesis of a PCR product that is an efficient substrate for recombination with a Donor (attP) Plasmid in the presence of GATEWAY™ BP Clonase™ Enzyme Mix. This reaction produces an Entry Clone of the PCR product (See Figure 8).

The conditions of the Gateward Cloning reaction with an attB PCR substrate are similar to those of the BP Reaction (see Example 10 below), except that the attB-PCR product (see Example 8) substitutes for the Expression Clone, and the attB-PCR positive control (attB-tet^r) substitutes for the Expression Clone Positive Control (GFP).

Materials needed:

- 5 X BP Reaction Buffer
- Desired attB-PCR product DNA, 50-100 ng in ≤ 8 µl TE.
- Donor (attP) Plasmid (Figures 49-54), 75 ng/µl, supercoiled DNA
- attB-tet^r PCR product positive control, 25 ng/µl
- GATEWAY™ BP Clonase™ Enzyme Mix (stored at - 80° C)
- 10x Clonase Stop Solution
- pUC19 DNA, 10 pg/µl.
- Chemically competent E.coli cells (competence: ≥ 1×10⁷ CFU/µg), 400 µl

Notes:

- Preparation of attB-PCR DNA: see Example 8.
- The Positive Control attB-tet^r PCR product contains a functional copy of the tet^r gene of pBR322, with its own promoter. By plating the transformation of the control BP Reaction on kanamycin (50 µg/ml) plates (if kan^r Donor Plasmids are used; see Figures 49-52) or an alternative selection agent (e.g., gentamycin, if gen^r Donor Plasmids are used; see Figure 54), and then picking about 50 of these colonies onto plates with tetracycline (20 µg/ml), the

-114-

percentage of Entry Clones containing functional tet^r among the colonies from the positive control reaction can be determined (% Expression Clones = (number of tet^r + kan^r (or gen^r) colonies/kan^r (or gen^r) colonies).

5

Procedure:

1. Assemble reactions as follows. Combine all components except GATEWAY™ BP Clonase™ Enzyme Mix, before removing GATEWAY™ BP Clonase™ Enzyme Mix from frozen storage.

10

Component	Neg.	Pos.	Test
	Tube 1	Tube 2	Tube 3
attB-PCR product, 50-100 ng			1 - 8 µl
Donor (attP) Plasmid 75 ng/µl	2 µl	2 µl	2 µl
attB-PCR tet ^r control DNA (75 ng/µl)		4 µl	
5 X BP Reaction Buffer	4 µl	4 µl	4 µl
TE	10 µl	6 µl	To 16 µl
GATEWAY™ BP Clonase™ Enzyme Mix (store at -80° C, add last)	4 µl	4 µl	4 µl
Total Volume	20 µl	20 µl	20 µl

15

20

25

30

2. Remove the GATEWAY™ BP Clonase™ Enzyme Mix from the -80° C freezer, place immediately on ice. The Clonase takes only a few minutes to thaw.
3. Add 4 µl of GATEWAY™ BP Clonase™ Enzyme Mix to the subcloning reaction, mix.
4. Return GATEWAY™ BP Clonase™ Enzyme Mix to - 80° C freezer.
5. Incubate tubes at 25° for at least 60 minutes.

-115-

6. Add 2 μ l Proteinase K (2 μ g/ μ l) to all reactions. Incubate for 20 min at 37°C.
7. Transform 2 μ l into 100 μ l competent E. coli, as per 3.2, above. Select on LB plates containing kanamycin, 50 μ g/ml.

5

Results:

10

15

20

25

30

In initial experiments, primers for amplifying tetR and ampR from pBR322 were constructed containing only the tetR- or ampR-specific targeting sequences, the targeting sequences plus attB1 (for forward primers) or attB2 (for reverse primers) sequences shown in Figure 9, or the attB1 or attB2 sequences with a 5' tail of four guanines. The construction of these primers is depicted in Figure 65. After PCR amplification of tetR and ampR from pBR322 using these primers and cloning the PCR products into host cells using the recombinational cloning system of the invention, the results shown in Figure 66 were obtained. These results demonstrated that primers containing attB sequences provided for a somewhat higher number of colonies on the tetracycline and ampicillin plates. However, inclusion of the 5' extensions of four or five guanines on the primers in addition to the attB sequences provided significantly better cloning results, as shown in Figures 66 and 67. These results indicate that the optimal primers for cloning of PCR products using recombinational cloning will contain the recombination site sequences with a 5' extension of four or five guanine bases.

To determine the optimal stoichiometry between attB-containing PCR products and attP-containing Donor plasmid, experiments were conducted where the amount of PCR product and Donor plasmid were varied during the BP Reaction. Reaction mixtures were then transformed into host cells and plated on tetracycline plates as above. Results are shown in Figure 68. These results indicate that, for optimal recombinational cloning results with a PCR product in the size range of the tet gene, the amounts of attP-containing Donor plasmids are between about 100-500 ng (most preferably about 200-300 ng), while the optimal concentrations of attB-containing PCR products is about 25-100 ng (most preferably about 100 ng), per 20 μ l reaction.

Experiments were then conducted to examine the effect of PCR product size on efficiency of cloning via the recombinational cloning approach of the invention.

-116-

PCR products containing attB1 and attB2 sites, at sizes 256 bp, 1 kb, 1.4 kb, 3.4 kb, 4.6 kb, 6.9 kb and 10.1 kb were prepared and cloned into Entry vectors as described above, and host cells were transformed with the Entry vectors containing the cloned PCR products. For each PCR product, cloning efficiency was calculated relative to cloning of pUC19 positive control plasmids as follows:

$$\text{Cloning Efficiency} = \frac{\text{CFU/ng attB PCR product}}{\text{CFU/ng pUC19 control}} \times \frac{\text{Size (kb) PCR product}}{\text{Size (kb) pUC19 control}}$$

5

10

15

The results of these experiments are depicted in Figures 69A-69C (for 256 bp PCR fragments), 70A-70C (for 1 kb PCR fragments), 71A-71C (for 1.4 kb PCR fragments), 72A-72C (for 3.4 kb PCR fragments), 73A-73C (for 4.6 kb PCR fragments), 74 (for 6.9 kb PCR fragments), and 75-76 (for 10.1 kb PCR fragments). The results shown in these figures are summarized in Figure 77, for different weights and moles of input PCR DNA.

20

25

Together, these results demonstrate that attB-containing PCR products ranging in size from about 0.25 kb to about 5 kb clone relatively efficiently in the recombinational cloning system of the invention. While PCR products larger than about 5 kb clone less efficiently (apparently due to slow resolution of cointegrates), longer incubation times during the recombination reaction appears to improve the efficiency of cloning of these larger PCR fragments. Alternatively, it may also be possible to improve efficiency of cloning of large (> about 5 kb) PCR fragments by using lower levels of input attP Donor plasmid and perhaps attB-containing PCR product, and/or by adjusting reaction conditions (e.g., buffer conditions) to favor more rapid resolution of the cointegrates.

Example 10: The BP Reaction

30

One purpose of the Gateward ("Entry") reaction is to convert an Expression Clone into an Entry Clone. This is useful when you have isolated an individual Expression Clone from an Expression Clone cDNA library, and you wish to transfer the nucleic acid molecule of interest into another Expression Vector, or

-117-

to move a population of molecules from an attB or attL library. Alternatively, you may have mutated an Expression Clone and now wish to transfer the mutated nucleic acid molecule of interest into one or more new Expression Vectors. In both cases, it is necessary first to convert the nucleic acid molecule of interest to an Entry Clone.

5

Materials needed:

- 5 X BP Reaction Buffer
- Expression Clone DNA, 100-300 ng in ≤ 8 µl TE.
- Donor (attP) Vector, 75 ng/µl, supercoiled DNA
- Positive control attB-tet-PCR DNA, 25 ng/µl
- GATEWAY™ BP Clonase™ Enzyme Mix (stored at - 80°C)
- Clonase Stop Solution (Proteinase K, 2 µg/µl).

10

15 Notes:

Preparation of the Expression Clone DNA: Miniprep DNA treated with RNase works well.

- 20 1. As with the LR Reaction (see Example 14), the BP Reaction is strongly influenced by the topology of the reacting DNAs. In general, the reaction is most efficient when one of the DNAs is linear and the other is supercoiled, compared to reactions where the DNAs are both linear or both supercoiled. Further, linearizing the attB Expression Clone (anywhere within the vector) will usually give more colonies than linearizing the Donor (attP) Plasmid. If finding a suitable cleavage site within your Expression Clone vector proves difficult, you may linearize the Donor (attP) Plasmid between the attP1 and attP2 sites (for example, at the *Nco*I site), avoiding the *ccdB* gene. Maps of Donor (attP) Plasmids are given in Figures 49-54.

25

30 Procedure:

1. Assemble reactions as follows. Combine all components at room temperature, except GATEWAY™ BP Clonase™ Enzyme Mix, before removing GATEWAY™ BP Clonase™ Enzyme Mix from freezer.

Component	Neg.	Pos.	Test
	Tube 1	Tube 2	Tube 3
Positive Control, attB-tet-PCR DNA, 25 ng/ μ l	4 μ l	4 μ l	
Desired attB Expression Clone DNA (100ng) linearized			1 - 8 μ l
Donor (attP) Plasmid, 75 ng/ μ l	2 μ l	2 μ l	2 μ l
5 X BP Reaction Buffer	4 μ l	4 μ l	4 μ l
TE	10 μ l	6 μ l	To 16 μ l
GATEWAY™ BP Clonase™ Enzyme Mix (store at - 80° C, add last)	---	4 μ l	4 μ l
Total Volume	20 μ l	20 μ l	20 μ l

- 5
2. Remove the GATEWAY™ BP Clonase™ Enzyme Mix from the -80°C freezer, place immediately on ice. The mixture takes only a few minutes to thaw.
 - 10 3. Add 4 μ l of GATEWAY™ BP Clonase™ Enzyme Mix to the subcloning reaction, mix.
 - 15 4. Return GATEWAY™ BP Clonase™ Enzyme Mix to - 80° C freezer.
 - 20 5. Incubate tubes at 25° for at least 60 minutes. If both the attB and attP DNAs are supercoiled, incubation for 2-24 hours at 25°C is recommended.
 - 25 6. Add 2 μ l Clonase Stop Solution. Incubate for 10 min at 37°C.
 - 30 7. Transform 2 μ l into 100 μ l competent E. coli, as above. Select on LB plates containing 50 μ g/ml kanamycin.

Example 11: Cloning PCR Products into Entry Vectors using Standard Cloning Methods

Preparation of Entry Vectors for Cloning of PCR Products

35 All of the Entry Vectors of the invention contain the death gene ccdB as a stuffer between the "left" and "right" restriction sites. The advantage of this arrangement is that there is virtually no background from vector that has not been cut with both restriction enzymes, because the presence of the ccdB gene will kill

all standard *E. coli* strains. Thus it is necessary to cut each Entry Vector twice, to remove the *ccdB* fragment.

We strongly recommend that, after digestion of the Entry Vector with the second restriction enzyme, you treat the reaction with phosphatase (calf intestine alkaline phosphatase, CIAP or thermosensitive alkaline phosphatase, TSAP). The phosphatase can be added directly to the reaction mixture, incubated for an additional time, and inactivated. This step dephosphorylates both the vector and *ccdB* fragments, so that during subsequent ligation there is less competition between the *ccdB* fragment and the DNA of interest for the termini of the Entry

10 Vector.

Blunt Cloning of PCR products

Generally PCR products do not have 5' phosphates (because the primers are usually 5' OH), and they are not necessarily blunt. (On this latter point, see Brownstein, et al., *BioTechniques* 20: 1006, 1996 for a discussion of how the sequence of the primers affects the addition of single 3' bases.) The following protocol repairs these two defects.

In a 0.5 ml tube, ethanol precipitate about 40 ng of PCR product (as judged from an agarose gel).

1. Dissolve the precipitated DNA in 10 μ l comprising 1 μ l 10 mM rATP, 1 μ l mixed 2 mM dNTPs (i.e., 2 mM each dATP, dCTP, dTTP, and dGTP), 2 μ l 5x T4 polynucleotide kinase buffer (350 mM Tris HCl (pH7.6), 50 mM MgCl₂, 500mM KCl, 5 mM 2-mercaptoethanol) 10 units T4 polynucleotide kinase, 1 μ l T4 DNA polymerase, and water to 10 μ l.
2. Incubate the tube at 37° for 10 minutes, then at 65° for 15 minutes, cool, centrifuge briefly to bring any condensate to the tip of the tube.
3. Add 5 μ l of the PEG/MgCl₂ solution, mix and centrifuge at room temperature for 10 minutes. Discard supernatant.
4. Dissolve the invisible precipitate in 10 μ l containing 2 μ l 5x T4 DNA ligase buffer (Life Technologies, Inc.), 0.5 units T4 DNA ligase, and about 50 ng of blunt, phosphatase-treated Entry Vector.

-120-

- 5 5. Incubate at 25° for 1 hour, then 65° for 10 minutes. Add 90 µl TE, transform
10 µl into 50 - 100 µl competent E. coli cells.
- 10 6. Plate on kanamycin.

5 Note: In the above protocol, steps b-c simultaneously polish the ends of the PCR product (through the exonuclease and polymerase activities of T4 DNA polymerase) and phosphorylate the 5' ends (using T4 polynucleotide kinase). It is necessary to inactivate the kinase, so that the blunt, dephosphorylated vector in step e cannot self ligate. Step d (the PEG precipitation) removes all small molecules (primers, nucleotides), and has also been found to improve the yield of cloned PCR product by 50 fold.

Cloning PCR Products after Digestion with Restriction Enzymes

15 Efficient cloning of PCR products that have been digested with restriction enzymes includes three steps: inactivation of *Taq* DNA polymerase, efficient restriction enzyme cutting, and removal of small DNA fragments.

20 Inactivation of *Taq* DNA Polymerase: Carryover of *Taq* DNA polymerase and dNTPs into a RE digestion significantly reduces the success in cloning a PCR product (D. Fox et al., *FOCUS* 20(1):15, 1998), because *Taq* DNA polymerase can fill in sticky ends and add bases to blunt ends. Either TAQQUENCH™ (obtainable from Life Technologies, Inc.; Rockville, Maryland) or extraction with phenol can be used to inactivate the *Taq*.

25 Efficient Restriction Enzyme Cutting: Extra bases on the 5' end of each PCR primer help the RE cut near ends of PCR products. With the availability of cheap primers, adding 6 to 9 bases on the 5' sides of the restriction sites is a good investment to ensure that most of the ends are digested. Incubation of the DNA with a 5-fold excess of restriction enzyme for an hour or more helps ensure success.

30 Removal of Small Molecules before Ligation: Primers, nucleotides, primer dimers, and small fragments produced by the restriction enzyme digestion,

-121-

can all inhibit or compete with the desired ligation of the PCR product to the cloning vector. This protocol uses PEG precipitation to remove small molecules.

Protocol for cutting the ends of PCR products with restriction enzyme(s):

5

1. Inactivation of Taq DNA polymerase in the PCR product:

Option A: Extraction with Phenol

A1. Dilute the PCR reaction to 200 µl with TE. Add an equal volume of phenol:chloroform:isoamyl alcohol, vortex vigorously for 20 seconds, and centrifuge for 1 minute at room temperature. Discard the lower phase.

10

A2. Extract the phenol from the DNA and concentrate as follows. Add an equal volume of 2-butanol (colored red with "Oil Red O" from Aldrich, if desired), vortex briefly, centrifuge briefly at room temperature. Discard the upper butanol phase. Repeat the extraction with 2-butanol. This time the volume of the lower aqueous phase should decrease significantly. Discard the upper 2-butanol phase.

15

A3. Ethanol precipitate the DNA from the aqueous phase of the above extractions. Dissolve in a 200 µl of a suitable restriction enzyme (RE) buffer.

20

Option B: Inactivation with TaqQuench

25

B1. Ethanol precipitate an appropriate amount of PCR product (100 ng to 1 µg), dissolve in 200 µl of a suitable RE buffer.

B2. Add 2 µl TaqQuench.

30

2. Add 10 to 50 units of restriction enzyme and incubate for at least 1 hour. Ethanol precipitate if necessary to change buffers for digestion at the other end of the PCR product.

-122-

3. Add ½ volume of the PEG/MgCl₂ mix to the RE digestion. Mix well and immediately centrifuge at room temperature for 10 minutes. Discard the supernatant (pellet is usually invisible), centrifuge again for a few seconds, discard any remaining supernatant.

5

4. Dissolve the DNA in a suitable volume of TE (depending on the amount of PCR product in the original amplification reaction) and apply an aliquot to an agarose gel to confirm recovery. Apply to the same gel 20-100 ng of the appropriate Entry Vector that will be used for the cloning.

10

Example 12: Determining The Expected Size of the GATEWAY™ Cloning Reaction Products

15 If you have access to a software program that will electronically cut and splice sequences, you can create electronic clones to aid you in predicting the sizes and restriction patterns of GATEWAY™ Cloning System recombination products.

20 The cleavage and ligation steps performed by the enzyme Int in the GATEWAY™ Cloning System recombination reactions mimic a restriction enzyme cleavage that creates a 7-bp 5'-end overhang followed by a ligation step that reseals the ends of the daughter molecules. The recombination proteins present in the Clonase cocktails (see Example 19 below) recognize the 15 bp core sequence present within all four types of att sites (in addition to other flanking sequences characteristic of each of the different types of att sites).

25 By treating these sites in your software program as if they were restriction sites, you can cut and splice your Entry Clones with various Destination Vectors and obtain accurate maps and sequences of the expected results from your GATEWAY™ Cloning System reactions.

30

Example 13: Protein Expression

Brief Review of Protein Expression

Transcription: The most commonly used promoters in *E. coli* Expression Vectors are variants of the lac promoter, and these can be turned on by adding

-123-

IPTG to the growth medium. It is usually good to keep promoters off until expression is desired, so that the host cells are not made sick by the overabundance of some heterologous protein. This is reasonably easy in the case of the lac promoters used in *E. coli*. One needs to supply the *lac I* gene (or its more productive relative, the *lac I^q* gene) to make *lac* repressor protein, which binds near the promoter and keeps transcription levels low. Some Destination Vectors for *E. coli* expression carry their own *lacI^q* gene for this purpose. (However, lac promoters are always a little "on," even in the absence of IPTG.)

Controlling transcription in eukaryotic cells is not nearly so straightforward or efficient. The tetracycline system of Bujard and colleagues is the most successful approach, and one of the Destination Vectors (pDEST11; Figure 31) has been constructed to supply this function.

Translation: Ribosomes convert the information present in mRNA into protein. Ribosomes scan RNA molecules looking for methionine (AUG) codons, which begin nearly all nascent proteins. Ribosomes must, however, be able to distinguish between AUG codons that code for methionine in the middle of proteins from those at the start. Most often ribosomes choose AUGs that are 1) first in the RNA (toward the 5' end), and 2) have the proper sequence context. In *E. coli* the favored context (first recognized by Shine and Dalgarno, *Eur. J. Biochem.* 57: 221 (1975)) is a run of purines (As and Gs) from five to 12 bases upstream of the initiating AUG, especially AGGAGG or some variant.

In eukaryotes, a survey of translated mRNAs by Kozak (*J. Biol. Chem.* 266: 19867 (1991)) has revealed a preferred sequence context, gcc Acc ATGG, around the initiating methionine, with the A at -3 being most important, and a purine at +4 (where the A of the ATG is +1), preferably a G, being next most influential. Having an A at -3 is enough to make most ribosomes choose the first AUG of an mRNA, in plants, insects, yeast, and mammals. (For a review of initiation of protein synthesis in eukaryotic cells, see: Pain, V.M. *Eur.J. Biochem.* 236:747-771, 1996.)

Consequences of Translation Signals for GATEWAY™ Cloning System: First, translation signals (Shine-Dalgarno in *E. coli*, Kozak in eukaryotes) have to be close to the initiating ATG. The attB site is 25 base pairs long. Thus if

translation signals are desired near the natural ATG of the nucleic acid molecule of interest, they must be present in the Entry Clone of that nucleic acid molecule of interest. Also, when a nucleic acid molecule of interest is moved from an Entry Clone to a Destination vector, any translation signals will move along. The result is that the presence or absence of Shine-Dalgarno and/or Kozak sequences in the Entry Clone must be considered, with the eventual Destination Vectors to be used in mind.

Second, although ribosomes choose the 5' ATG most often, internal ATGs are also used to begin protein synthesis. The better the translation context around this internal ATG, the more internal translation initiation will be seen. This is important in the GATEWAY™ Cloning System, because you can make an Entry Clone of your nucleic acid molecule of interest, and arrange to have Shine-Dalgarno and/or Kozak sequences near the ATG. When this cassette is recombined into a Destination Vector that transcribes your nucleic acid molecule of interest, you get native protein. If you want, you can make a fusion protein in a different Destination Vector, since the Shine-Dalgarno and/or Kozak sequences do not contain any stop signals in the same reading frame. However, the presence of these internal translation signals may result in a significant amount of native protein being made, contaminating, and lowering the yield of, your fusion protein.

This is especially likely with short fusion tags, like His6.

A good compromise can be recommended. If an Entry Vector like pENTR7 (Figure 16) or pENTR8 (Figure 17) is chosen, the Kozak bases are present for native eukaryotic expression. The context for *E. coli* translation is poor, so the yield of an amino-terminal fusion should be good, and the fusion protein can be digested with the TEV protease to make near-native protein following purification.

Recommended Conditions for Synthesis of Proteins in E. coli: When making proteins in *E. coli* it is advisable, at least initially, to incubate your cultures at 30°C, instead of at 37°C. Our experience indicates that proteins are less likely to form aggregates at 30°C. In addition, the yields of proteins from cells grown at 30°C frequently are improved.

The yields of proteins that are difficult to express may also be improved by inducing the cultures in mid-log phase of growth, using cultures begun in the morning from overnight growths, as opposed to harvesting directly from an overnight culture. In the latter case, the cells are preferably in late log or stationary growth, which can favor the formation of insoluble aggregates.

5

Example 14: Constructing Destination Vectors from Existing Vectors

Destination Vectors function because they have two recombination sites, attR1 and attR2, flanking a chloramphenicol resistance (CmR) gene and a death gene, ccdB. The GATEWAY™ Cloning System recombination reactions exchange the entire Cassette (except for a few bases comprising part of the attB sites) for the DNA segment of interest from the Entry Vector. Because attR1, CmR, ccdB gene, and attR2 are contiguous, they can be moved on a single DNA segment. If this Cassette is cloned into a plasmid, the plasmid becomes a Destination Vector. Figure 63 shows a schematic of the GATEWAY™ Cloning System Cassette; attR cassettes in all three reading frames contained in vectors pEZC15101, pEZC15102 and pEZC15103 are shown in Figures 64A, 64B, and 64C, respectively.

10

15

20

25

30

The protocol for constructing a Destination Vector is presented below.

Keep in mind the following points:

- Destination Vectors must be constructed and propagated in one of the DB strains of *E. coli* (e.g., DB3.1, and particularly *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells) available from Life Technologies, Inc. (and described in detail in U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, which is incorporated herein by reference), because the ccdB death gene will kill any *E. coli* strain that has not been mutated such that it will survive the presence of the ccdB gene.
- If your Destination Vector will be used to make a fusion protein, a GATEWAY™ Cloning System cassette with the correct reading frame must be used. The nucleotide sequences of the ends of the cassettes are shown in Figure 78. The reading frame of the fusion protein domain must

-126-

be in frame with the core region of the attR1 site (for an amino terminal fusion) so that the six As are translated into two lysine codons. For a C-terminal fusion protein, translation through the core region of the attR2 site should be in frame with -TAC-AAA-, to yield -Tyr-Lys-.

- 5 • Note that each reading frame Cassette has a different unique restriction site between the chloramphenicol resistance and ccdB genes (*Mlu*I for reading frame A, *Bgl*II for reading frame B, and *Xba*I for reading frame C; see Figure 63).
- 10 • Most standard vectors can be converted to Destination Vectors, by inserting the Entry Cassette into the MCS of that vector.

Protocol for Making a Destination Vector

15 1. If the vector will make an amino fusion protein, it is necessary to keep the "aaa aaa" triplets in attR1 in phase with the triplets of the fusion protein. Determine which Entry cassette to use as follows:

20 a.) Write out the nucleotide sequence of the existing vector near the restriction site into which the Entry cassette will be cloned. These must be written in triplets corresponding to the amino acid sequence of the fusion domain.

25 b.) Draw a vertical line through the sequence that corresponds to the restriction site end, after it has been cut and made blunt, i.e., after filling in a protruding 5' end or polishing a protruding 3' end.

30 c.) Choose the appropriate reading frame cassette:

- If the coding sequence of the blunt end ends after a complete codon triplet, use the reading frame A cassette. See Figures 78, 79 and 80.

-127-

• If the coding sequence of the blunt end ends in a single base, use the reading frame B cassette. See Figures 78, 79 and 81.

5 • If the coding sequence of the blunt end ends in two bases, use the reading frame C cassette. See Figures 78, 79, 82A-B, and 83A-C.

10 2. Cut one to five micrograms of the existing plasmid at the position where you wish your nucleic acid molecule of interest (flanked by att sites) to be after the recombination reactions. Note: it is better to remove as many of the MCS restriction sites as possible at this step. This makes it more likely that restriction enzyme sites within the GATEWAY™ Cloning System Cassette will be unique in the new plasmid, which is important for linearizing the Destination Vector (Example 14, below).

15 3. Remove the 5' phosphates with alkaline phosphatase. While this is not mandatory, it increases the probability of success.

20 4. Make the end(s) blunt with fill-in or polishing reactions. For example, to 1 µg of restriction enzyme-cut, ethanol-precipitated vector DNA, add:

- i. 20 µl 5x T4 DNA Polymerase Buffer (165 mM Tris-acetate (pH 7.9), 330 mM Na acetate, 50 mM Mg acetate, 500 µg/ml BSA, 2.5 mM DTT)
- ii. 5 µl 10mM dNTP mix
- iii. 1 Unit of T4 DNA Polymerase
- iv. Water to a final volume of 100 µl
- v. Incubate for 15 min at 37°C.

25 5. Remove dNTPs and small DNA fragments: Ethanol precipitate (add three volumes of room temperature ethanol containing 0.1 M sodium acetate, mix well, immediately centrifuge at room temperature 5 - 10 minutes), dissolve wet precipitate in 200 µl TE, add 100 µl 30% PEG 8000, 30 mM MgCl₂, mix well,

-128-

immediately centrifuge for 10 minutes at room temperature, discard supernatant, centrifuge again a few seconds, discard any residual liquid.

6. Dissolve the DNA to a final concentration of 10 - 50 ng per microliter. Apply 5 20 - 100 ng to a gel next to supercoiled plasmid and linear size standards to confirm cutting and recovery. The cutting does not have to be 100% complete, since you will be selecting for the chloramphenicol marker on the Entry cassette.

10 7. In a 10 μ l ligation reaction combine 10 - 50 ng vector, 10 - 20 ng of Entry Cassette (Figure 79), and 0.5 units T4 DNA ligase in ligase buffer. After one hour (or overnight, whichever is most convenient), transform 1 μ l into one of the DB strains of competent *E. coli* cells with a *gyrA*462 mutation (See U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, which is incorporated herein by reference), preferably DB3.1, and most preferably *E. coli* LIBRARY 15 EFFICIENCY[®] DB3.1TM Competent Cells. The *ccdB* gene on the Entry Cassette will kill other strains of *E. coli* that have not been mutated so as to survive the presence of the *ccdB* gene.

20 8. After expression in SOC medium, plate 10 μ l and 100 μ l on chloramphenicol-containing (30 μ g / ml) plates, incubate at 37° C.

25 9. Pick colonies, make miniprep DNA. Treat the miniprep with RNase A and store in TE. Cut with the appropriate restriction enzyme to determine the orientation of the Cassette. Choose clones with the attR1 site next to the amino end of the protein expression function of the plasmid.

Notes on Using Destination Vectors

- We have found that about ten-fold more colonies result from a GATEWAY[™] Cloning System reaction if the Destination Vector is linear or relaxed. If the 30 competent cells you use are highly competent ($>10^8$ per microgram), linearizing the Destination Vector is less essential.

-129-

- The site or sites used for the linearization must be within the Entry Cassette. Sites that cut once or twice within each cassette are shown in Figures 80-82.
- Minipreps of Destination Vectors will work fine, so long as they have been treated with RNase. Since most DB strains are *endA*- (See U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, which is incorporated herein by reference), minipreps can be digested with restriction enzymes without a prior phenol extraction.
- Reading the OD₂₆₀ of miniprep DNA is inaccurate unless the RNA and ribonucleotides have been removed, for example, by a PEG precipitation.

5

10

Example 15: Some Options in Choosing Appropriate Entry Vectors and Destination Vectors: An Example

15

In some applications, it may be desirable to express a nucleic acid molecule of interest in two forms: as an amino-terminal fusion in *E. coli*, and as a native protein in eukaryotic cells. This may be accomplished in any of several ways:

20

Option 1: Your choices depend on your nucleic acid molecule of interest and the fragment that contains it, as well as the available Entry Vectors. For eukaryotic translation, you need consensus bases according to Kozak (*J. Biol. Chem.* 266:19867, 1991) near the initiating methionine (ATG) codon. All of the Entry Vectors offer this motif upstream of the *XmnI* site (blunt cutter). One option is to amplify your nucleic acid molecule of interest, with its ATG, by PCR, making the amino end blunt and the carboxy end containing the natural stop codon followed by one of the "right side" restriction sites (*EcoRI*, *NorI*, *XbaI*, *EcoRV*, or *XbaI* of the pENTR vectors).

25

If you know your nucleic acid molecule of interest does not have, for example, an *XbaI* site, you can make a PCR product that has this structure:

30

Xba I

5' ATG nnn nnn --- nnn TAA ctc gag nnn nnn 3'
3' tac nnn nnn --- nnn att gag ctc nnn nnn 5'

-130-

After cutting with *XhoI*, the fragment is ready to clone:

5' ATG nnn nnn --- nnn TAA c 3'

3' tac nnn nnn --- nnn att gag ct 5'

5 (If you follow this example, don't forget to put a phosphate on the amino oligo.)

10

Option 2: This PCR product could be cloned into two Entry Vectors to give the desired products, between the *XmnI* and *XhoI* sites: pENTR1A (Figures 10A, 10B) or pENTR7 (Figures 16A, 16B). If you clone into pENTR1A, amino fusions will have the minimal number of amino acids between the fusion domain and your nucleic acid molecule of interest, but the fusion cannot be removed with TEV protease. The converse is true of clones in pENTR7, i.e., an amino fusion can be cleaved with TEV protease, at the cost of more amino acids between the fusion and your nucleic acid molecule of interest.

15

In this example, let us choose to clone our hypothetical nucleic acid molecule of interest into pENTR7, between the *XmnI* and *XhoI* sites. Once this is accomplished, several optional protocols using the Entry Clone pENTR7 may be followed:

20

Option 3: Since the nucleic acid molecule of interest has been amplified with PCR, it may be desirable to sequence it. To do this, transfer the nucleic acid molecule of interest from the Entry Vector into a vector that has priming sites for the standard sequencing primers. Such a vector is pDEST6 (Figures 26A, 26B). This Destination Vector places the nucleic acid molecule of interest in the opposite orientation to the lac promoter (which is leaky -- see Example 3 above). If the gene product is toxic to *E. coli*, this Destination Vector will minimize its toxicity.

25

30

Option 4: While the sequencing is going on, you might wish to check the expression of the nucleic acid molecule of interest in, for example, CHO cells, by recombining the nucleic acid molecule of interest into a CMV promoter vector (pDEST7, Figure 27; or pDEST12, Figure 32), or into a baculovirus vector (pDEST8, Figure 28; or pDEST10, Figure 30) for expression in insect cells. Both

-131-

of these vectors will transcribe the coding sequence of your nucleic acid molecule of interest, and translate it from the ATG of the PCR product using the Kozak bases upstream of the *XmnI* site.

5 **Option 5:** If you wish to purify protein, for example to make antibodies, you can clone the nucleic acid molecule of interest into a His6 fusion vector, pDEST2 (Figure 22). Since the nucleic acid molecule of interest is cloned downstream of the TEV protease cleavage domain of pENTR7 (Figure 16), the amino acid sequence of the protein produced will be:

10 [----- attB1 -----] TEV protease
NH2- MSYYHHHHHGGITSLYKKAGFENLYFQ! GTM---COOH

15 The attB site and the restriction sites used to make the Destination and Entry Vectors are translated into the underlined 11 amino acids (GITSLYKKAGF). Cleavage with TEV protease (arrow) leaves two amino acids, GT, on the amino end of the gene product.

20 See Figure 55 for an example of a nucleic acid molecule of interest, the chloramphenicol acetyl transferase (CAT) gene, cloned into pENTR7 (Figure 16) as a blunt (amino)-*Xba*I (carboxy) fragment, then cloned by recombination into the His6 fusion vector pDEST2 (Figure 22).

25 **Option 6:** If the His6 fusion protein is insoluble, you may go on and try a GST fusion. The appropriate Destination vector is pDEST3 (Figure 23).

30 **Option 7:** If you need to make RNA probes and prefer SP6 RNA polymerase, you can make the top strand RNA with your nucleic acid molecule of interest, cloned into pSPORT+ (pDEST5 (Figures 25A, 25B)), and the bottom strand RNA with the nucleic acid molecule of interest cloned into pSPORT(-) (pDEST6 (Figures 26A, 26B)). Opposing promoters for T7 RNA polymerase and SP6 RNA polymerase are also present in these clones.

-132-

5 **Option 8:** It is often worthwhile to clone your nucleic acid molecule of interest into a variety of Destination Vectors in the same experiment. For example, if the number of colonies varies widely when the various recombination reactions are transformed into *E. coli*, this may be an indication that the nucleic acid molecule of interest is toxic in some contexts. (This problem is more clearly evident when a positive control gene is used for each Destination Vector.) Specifically, if many more colonies are obtained when the nucleic acid molecule of interest is recombined into pDEST6 than in pDEST5, there is a good chance that leakiness of the lac promoter is causing some expression of the nucleic acid molecule of interest in pSPORT "+" (which is not harmful in pDEST6 because the nucleic acid molecule of interest is in the opposite orientation).

10

15 ***Example 16: Demonstration of a One-tube Transfer of a PCR Product (or Expression Clone) to Expression Clone via a Recombinational Cloning Reaction***

20 In the BxP recombination (Entry or Gateward) reaction described herein, a DNA segment flanked by attB1 and attB2 sites in a plasmid conferring ampicillin resistance was transferred by recombination into an attP plasmid conferring kanamycin resistance, which resulted in a product molecule wherein the DNA segment was flanked by attL sites (attL1 and attL2). This product plasmid comprises an "attL Entry Clone" molecule, because it can react with a "attR Destination Vector" molecule via the LxR (Destination) reaction, resulting in the transfer of the DNA segment to a new (ampicillin resistant) vector. In the previously described examples, it was necessary to transform the BxP reaction products into *E. coli*, select kanamycin resistant colonies, grow those colonies in liquid culture, and prepare miniprep DNA, before reacting this DNA with a Destination Vector in an LxR reaction.

25

30 The goal of the following experiment was to eliminate the transformation and miniprep DNA steps, by adding the BxP Reaction products directly to an LxR Reaction. This is especially appropriate when the DNA segment flanked by attB sites is a PCR product instead of a plasmid, because the PCR product cannot give

-133-

ampicillin-resistant colonies upon transformation, whereas attB plasmids (in general) carry an ampicillin resistance gene. Thus use of a PCR product flanked by attB sites in a BxP Reaction allows one to select for the ampicillin resistance encoded by the desired attB product of a subsequent LxR Reaction.

5 Two reactions were prepared: Reaction A, negative control, no attB PCR product, (8 µl) contained 50 ng pEZC7102 (attP Donor plasmid, confers kanamycin resistance) and 2 µl BxP Clonase (22 ng / µl Int protein and 8 ng/µl IHF protein) in BxP buffer (25 mM Tris HCl, pH 7.8, 70 mM KCl, 5 mM spermidine, 0.5 mM EDTA, 250 µg / ml BSA). Reaction B (24 µl) contained
10 150 ng pEZC7102, 6 µl BxP Clonase, and 120 ng of the attB -tet-PCR product in the same buffer as reaction A. The attB - tet - PCR product comprised the tetracycline resistance gene of plasmid pBR322, amplified with two primers containing either attB1 or attB2 sites, and having 4 Gs at their 5' ends, as described earlier.

15 The two reactions were incubated at 25°C for 30 minutes. Then aliquots of these reactions were added to new components that comprised LxR Reactions or appropriate controls for the LxR Reaction. Five new reactions were thus produced:

20 **Reaction 1:** 5 µl of reaction A was added to a 5 µl LxR Reaction containing 25 ng NcoI-cut pEZC8402 (the attR Destination Vector plasmid) in LxR buffer (37.5 mM Tris HCl, pH 7.7, 16.5 mM NaCl, 35 mM KCl, 5 mM spermidine, 375 µg / ml BSA), and 1 µl of GATEWAY™ LR Clonase™ Enzyme Mix (total volume of 10 µl).

25 **Reaction 2:** Same as reaction 1, except 5 µl of reaction B (positive) were added instead of reaction A (negative).

30 **Reaction 3:** Same as reaction 2, except that the amounts of Nco-cut pEZC8402 and GATEWAY™ LR Clonase™ Enzyme Mix were doubled, to 50 ng and 2 µl, respectively.

-134-

Reaction 4: Same as reaction 2, except that 25 ng of pEZ11104 (a positive control attL Entry Clone plasmid) were added in addition to the aliquot of reaction B.

Reaction 5: Positive control LxR Reaction, containing 25 ng *Nco*I-cut pEZZC8402, 25 ng pEZ11104, 37.5 mM Tris HCl pH 7.7, 16.5 mM NaCl, 35 mM KCl, 5 mM spermidine, 375 µg / ml BSA and 1 µl GATEWAY™ LR Clonase™ Enzyme Mix in a total volume of 5 µl.

All five reactions were incubated at 25°C for 30 minutes. Then, 1 µl aliquots of each of the above five reactions, plus 1 µl from the remaining volume of Reaction B, the standard BxP Reaction, were used to transform 50 µl competent DH5 α *E. coli*. DNA and cells were incubated on ice for 15 min., heat shocked at 42°C for 45 sec., and 450 µl SOC were added. Each tube was incubated with shaking at 37°C for 60 min. Aliquots of 100 µl and 400 µl of each transformation were plated on LB plates containing either 50 µg/ml kanamycin or 100 µg/ml ampicillin (see Table 2). A transformation with 10 pg of pUC19 DNA (plated on LB-amp₁₀₀) served as a control on the transformation efficiency of the DH5 α cells. Following incubation overnight at 37°C, the number of colonies on each plate was determined.

Results of these reactions are shown in Table 2.

Table 2*

Reaction No.:	1	2	3	4	5	6
Number of Colonies						
Vol. plated:	Neg. Control BxP Reaction	1X pEZZC8402 and LR Clonase™	2X pEZZC8402 and LR Clonase™	LxR Reaction with Pos. Control DNA	LxR Reaction alone	BxP Reaction alone
100 µl	2	1	8	9	~1000	~1000
400 µl	5	10	35	62	>2000	>2000
Selection:	Kan	Amp	Amp	Amp	Amp	Kan

*(Transformation with pUC 19 DNA yielded 1.4 x 10⁹ CFU/µg DNA.)

-135-

34 of the 43 colonies obtained from Reaction 3 were picked into 2 ml Terrific Broth with 100 µg/ml ampicillin and these cultures were grown overnight, with shaking, at 37°C. 27 of the 34 cultures gave at least moderate growth, and of these 24 were used to prepare miniprep DNA, using the standard protocol.

5 These 24 DNAs were initially analyzed as supercoiled (SC) DNA on a 1% agarose gel to identify those with inserts and to estimate the sizes of the inserts. Fifteen of the 24 samples displayed SC DNA of the size predicted (5553 bp) if *tetx7102* had correctly recombined with pEZC8402 to yield *tetx8402*. One of these samples contained two plasmids, one of ~5500 bp and a one of ~3500 bp. The majority of the remaining clones were approximately 4100 bp in size.

10

All 15 of the clones displaying SC DNA of predicted size (~5500 bp) were analyzed by two different double digests with restriction endonucleases to confirm the structure of the expected product: *tetx8402*. (See plasmid maps, Figures 57-59) In one set of digests, the DNAs were treated with *Not I* and *Eco RI*, which should cut the predicted product just outside both *attB* sites, releasing the *tet^r* insert on a fragment of 1475 bp. In the second set of digests, the DNAs were digested with *NoI* and with *NruI*. *NruI* cleaves asymmetrically within the subcloned *tet^r* insert, and together with *NoI* will release a fragment of 1019 bp.

15

Of the 15 clones analyzed by double restriction digestion, 14 revealed the predicted sizes of fragments for the expected product.

20

Interpretation:

The DNA components of Reaction B, pEZC7102 and *attB-tet-PCR*, are shown in Figure 56. The desired product of BxP Reaction B is *tetx7102*, depicted in Figure 57. The LxR Reaction recombines the product of the BxP Reaction, *tetx7102* (Figure 57), with the Destination Vector, pEZC8402, shown in Figure 58. The LxR Reaction with *tetx7102* plus pEZC8402 is predicted to yield the desired product *tetx8402*, shown in Figure 59.

25

Reaction 2, which combined the BxP Reaction and LxR Reaction, gave few colonies beyond those of the negative control Reaction. In contrast, Reaction 3, with twice the amount of pEZC8402 (Figure 58) and LxR Clonase, yielded a

30

larger number of colonies. These colonies were analyzed further, by restriction digestion, to confirm the presence of expected product. Reaction 4 included a known amount of attL Entry Clone plasmid in the combined BxP-plus-LxR reaction. But reaction 4 yielded only about 1% of the colonies obtained when the same DNA was used in a LxR reaction alone, Reaction 6. This result suggests that the LxR reaction may be inhibited by components of the BxP reaction.

Restriction endonuclease analysis of the products of Reaction 3 revealed that a sizeable proportion of the colonies (14 of the 34 analyzed) contained the desired tet^r subclone, tetx8402 (Figure 59).

The above results establish the feasibility of performing first a BxP recombination reaction followed by a LxR recombination reaction -- in the same tube -- simply by adding the appropriate buffer mix, recombination proteins, and DNAs to a completed BxP reaction. This method should prove useful as a faster method to convert attB-containing PCR products into different Expression Clones, eliminating the need to isolate first the intermediate attL-PCR insert subclones, before recombining these with Destination Vectors. This may prove especially valuable for automated applications of these reactions.

This same one-tube approach allows for the rapid transfer of nucleic acid molecules contained in attB plasmid clones into new functional vectors as well. As in the above examples, attL subclones generated in a BxP Reaction can be recombined directly with various Destination Vectors in a LxR reaction. The only additional requirement for using attB plasmids, instead of attB-containing PCR products, is that the Destination Vector(s) employed must contain a different selection marker from the one present on the attB plasmid itself and the attP vector.

Two alternative protocols for a one-tube reaction have also proven useful and somewhat more optimal than the conditions described above.

Alternative 1:

Reaction buffer contained 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 0.25 mM EDTA, 2.5 mM spermidine, and 200 µg/ml BSA. After a 16 (or 3) hour incubation of the PCR product (100 ng) + attP Donor plasmid (100 ng) +

-137-

5

10

GATEWAY™ BP Clonase™ Enzyme Mix + Destination Vector (100 ng), 2 μ l of GATEWAY™ LR Clonase™ Enzyme Mix (per 10 μ l reaction mix) was added and the mixture was incubated an additional 6 (or 2) hours at 25°C. Stop solution was then added as above and the mixture was incubated at 37°C as above and transformed by electroporation with 1 μ l directly into electrocompetent host cells. Results of this series of experiments demonstrated that longer incubation times (16 hours vs. 3 hours for the BP Reaction, 6 hours vs. 2 hours for the LR Reaction) resulted in about twice as many colonies being obtained as for the shorter incubation times. With two independent genes, 10/10 colonies having the correct cloning patterns were obtained.

Alternative 2:

15

A standard BP Reaction under the reaction conditions described above for Alternative 1 was performed for 2 hours at 25°C. Following the BP Reaction, the following components were added to the reaction mixture in a total volume of 7 μ l:

20 mM Tris-HCl, pH 7.5

100 mM NaCl

5 μ g/ml Xis-His6

15% glycerol

~1000 ng of Destination Vector

20

25

30

The reaction mixture was then incubated for 2 hours at 25°C, and 2.5 μ l of stop solution (containing 2 μ g/ml proteinase K) was added and the mixture was incubated at 37°C for an additional 10 minutes. Chemically competent host cells were then transformed with 2 μ l of the reaction mixture, or electrocompetent host cells (e.g., EMax DH10B cells; Life Technologies, Inc.) were electroporated with 2 μ l of the reaction mixture per 25-40 μ l of cells. Following transformation, mixtures were diluted with SOC, incubated at 37°C, and plated as described above on media selecting for the selection markers on the Destination Vector and the Entry clone (B x P reaction product). Analogous results to those described for Alternative 1 were obtained with these reaction conditions -- a higher level of colonies containing correctly recombined reaction products were observed.

Example 17: Demonstration of a One-tube Transfer of a PCR Product (or Expression Clone) to Expression Clone via a Recombinational Cloning Reaction

5 Single-tube transfer of PCR product DNA or Expression Clones into Expression Clones by recombinational cloning has also been accomplished using a procedure modified from that described in Example 16. This procedure is as follows:

- 10 • Perform a standard BP (Gateward) Reaction (see Examples 9 and 10) in 20 µl volume at 25°C for 1 hour.

- 15 • After the incubation is over, take a 10 µl aliquot from the 20 µl total volume and add 1 µl of Proteinase K (2 mg/ml) and incubate at 37°C for 10 minutes. This first aliquot can be used for transformation and gel assay of BP reaction analysis. Plate BP reaction transformation on LB plates with Kanamycin (50 µg/ml).

- 20 • Add the following reagents to the remaining 10 µl aliquot of the BP reaction:

1 µl of 0.75 M NaCl

2 µl of destination vector (150 ng/µl)

4 µl of LR Clonase™ (after thawing and brief mixing)

- 25 • Mix all reagents well and incubate at 25°C for 3 hours. Stop the reaction at the end of incubation with 1.7 µl of Proteinase K (2 mg/ml) and incubate at 37°C for 10 minutes.

- 30 • Transform 2 µl of the completed reaction into 100 µl of competent cells. Plate 100 µl and 400 µl on LB plates with Ampicillin (100 µg/ml).

Notes:

- If your competent cells are less than 10⁸ CFU/µg, and you are concerned about getting enough colonies, you can improve the yield several fold by incubating the

-139-

BP reaction for 6-20 hours. Electroporation also can yield better colony output than chemical transformation.

•PCR products greater than about 5-6 kb show significantly lower cloning efficiency in the BP reaction. In this case, we recommend using longer incubation times for both BP and LR steps.

•If you want to move your insert gene into several destination vectors simultaneously, then scale up the initial BP reaction volume so that you have a 10 µl aliquot for adding each destination vector.

Example 18: Optimization of GATEWAY™ Clonase™ Enzyme Compositions

The enzyme compositions containing Int and IHF (for BP Reactions) were optimized using a standard functional recombinational cloning reaction (a BP reaction) between attB-containing plasmids and attP-containing plasmids, according to the following protocol:

Materials and Methods:

Substrates:

AttP - supercoiled pDONR201

AttB - linear ~ 1Kb [³H]PCR product amplified from pEYC7501

Proteins:

IntH6 -- His₆-carboxy- tagged λ Integrase

IHF -- Integration Host Factor

Clonase:

50 ng/µl IntH6 and 20 ng/µl IHF, admixed in 25 mM Tris- HCl (pH 7.5),

22 mM NaCl, 5 mM EDTA, 1 mg/ml BSA, 5 mM Spermidine, and 50% glycerol.

-140-

Reaction Mixture (total volume of 40 µl):

1000 ng AttP plasmid

600 ng AttB [³H] PCR product

8 µl Clonase (400 ng IntH6, 160 ng IHF) in 25 mM Tris-HCl (pH 7.5),

5 22 mM NaCl, 5 mM EDTA, 1 mg/ml BSA, 5 mM Spermidine, 5 mM DTT.

Reaction mixture was incubated for 1 hour at 25°C, 4 µl of 2 µg/µl proteinase K was added and mixture was incubated for an additional 20 minutes at 37°C. Mixture was then extracted with an equal volume of Phenol/Chloroform/Isoamyl alcohol. The aqueous layer was then collected, and 0.1 volumes of 3 M sodium acetate and 2 volumes of cold 100% ethanol were added. Tubes were then spun in a microcentrifuge at maximum RPM for 10 minutes at room temperature. Ethanol was decanted, and pellets were rinsed with 70% ethanol and re-centrifuged as above. Ethanol was decanted, and pellets were allowed to air dry for 5-10 minutes and then dissolved in 20 µl of 33 mM Tris-Acetate (pH 7.8), 66 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, and 1mM ATP. 2 units of exonuclease V (e.g., Plasmid Safe; EpiCentre, Inc., Madison, WI) was then added, and the mixture was incubated at 37°C for 30 minutes.

20 Samples were then TCA-washed by spotting 30 µl of reaction mixture onto a Whatman GF/C filter, washing filters once with 10% TCA + 1% NaPPi for 10 minutes, three times with 5% TCA for 5 minutes each, and twice with ethanol for 5 minutes each. Filters were then dried under a heat lamp, placed into a scintillation vial, and counted on a β liquid scintillation counter (LSC).

25 The principle behind this assay is that, after exonuclease V digestion, only double-stranded circular DNA survives in an acid-insoluble form. All DNA substrates and products that have free ends are digested to an acid-soluble form and are not retained on the filters. Therefore, only the ³H-labeled attB linear DNA which ends up in circular form after both inter- and intramolecular integration is complete is resistant to digestion and is recovered as acid-insoluble product. Optimal enzyme and buffer formulations in the Clonase compositions therefore are those that give the highest levels of circularized ³H-labeled attB-containing

-141-

sequences, as determined by highest cpm in the LSC. Although this assay was designed for optimization of GATEWAY™ BP Clonase™ Enzyme Mix compositions (Int + IHF), the same type of assay may be performed to optimize GATEWAY™ LR Clonase™ Enzyme Mix compositions (Int + IHF + Xis), except that the reaction mixtures would comprise 1000 ng of AttR (instead of AttP) and 600 ng of AttL (instead of AttB), and 40 ng of His₆-carboxy- tagged Xis (XisH6) in addition to the IntH6 and IHF.

Example 19: Testing Functionality of Entry and Destination Vectors

As part of assessment of the functionality of particular vectors of the invention, it is important to functionally test the ability of the vectors to recombine. This assessment can be carried out by performing a recombinational cloning reaction (as schematized in Figures 2, 4, and 5A and 5B, and as described herein and in commonly owned U.S. Application Nos. 08/486,139, filed June 7, 1995, 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed October 23, 1998, the disclosures of all of which are incorporated by reference herein in their entireties), by transforming E. coli and scoring colony forming units. However, an alternative assay may also be performed to allow faster, more simple assessment of the functionality of a given Entry or Destination Vector by agarose gel electrophoresis. The following is a description of such an in vitro assay.

Materials and Methods:

Plasmid templates pEYC1301 (Figure 84) and pEYC1313 (Figure 85), each containing a single wild type att site, were used for the generation of PCR products containing attL or attR sites, respectively. Plasmid templates were linearized with *Afl*NI, phenol extracted, ethanol precipitated and dissolved in TE to a concentration of 1 ng/μl.

-142-

PCR primers (capital letters represent base changes from wildtype):

attL1 gggg agcct gcttttGtacAaa gttggcatta taaaaaagca ttgc
attL2 gggg agcct gctttCttGtacAaa gttggcatta taaaaaagca ttgc
attL right tggccggg aagctagagt aa

5 attR1 gggg Acaag ttTgtCAaaaaaagc tgaacgaga aacgtaaaat
attR2 gggg Acaag ttTgtCaaGaaagc tgaacgaga aacgtaaaat
attR right ca gacggcatga tgaacctgaa

10 PCR primers were dissolved in TE to a concentration of 500 pmol/ μ l. Primer mixes were prepared, consisting of attL1 + attLright primers, attL2 + attLright primers, attR1 + attRright primers, and attR2 + attRright primers, each mix containing 20 pmol/ μ l of each primer.

PCR reactions:

15 1 μ l plasmid template (1 ng)
1 μ l primer pairs (20 pmoles of each)
3 μ l of H₂O
45 μ l of Platinum PCR SuperMix® (Life Technologies, Inc.)

20 Cycling conditions (performed in MJ thermocycler):

95 °C/2 minutes
94 °C/30 seconds
25 cycles of 58 °C/30 seconds and 72 °C/1.5 minutes
72 °C/5 minutes
25 5 °C/hold

The resulting attL PCR product was 1.5 kb, and the resulting attR PCR product was 1.0 kb.

30 PCR reactions were PEG/MgCl₂ precipitated by adding 150 μ l H₂O and 100 μ l of 3x PEG/ MgCl₂ solution followed by centrifugation. The PCR products were dissolved in 50 μ l of TE. Quantification of the PCR product was performed by gel electrophoresis of 1 μ l and was estimated to be 50-100 ng/ μ l.

Recombination reactions of PCR products containing attL or attR sites with GATEWAY™ plasmids was performed as follows:

8 µl of H₂O

2 µl of attL or attR PCR product (100-200 ng)

5 2 µl of GATEWAY™ plasmid (100 ng)

4 µl of 5x Destination buffer

4 µl of GATEWAY™ LR Clonase™ Enzyme Mix

20 µl total volume (the reactions can be scaled down to a 5 µl total volume by adjusting the volumes of the components to about 1/4 of those shown above, while keeping the stoichiometries the same).

10 Clonase reactions were incubated at 25 °C for 2 hours. 2 µl of proteinase K (2 mg/ml) was added to stop the reaction. 10 µl was then run on a 1 % agarose gel. Positive control reactions were performed by reacting attL1 PCR product (1.0 kb) with attR1 PCR product (1.5 kb) and by similarly reacting attL2 PCR product with attR2 PCR product to observe the formation of a larger (2.5 kb) recombination product. Negative controls were similarly performed by reacting attL1 PCR product with attR2 PCR product and vice versa or reactions of attL PCR product with an attL plasmid, etc.

15 In alternative assays, to test attB Entry vectors, plasmids containing single attP sites were used. Plasmids containing single att sites could also be used as recombination substrates in general to test all Entry and Destination vectors (*i.e.*, those containing attL, attR, attB and attP sites). This would eliminate the need to do PCR reactions.

25

Results:

30

Destination and Entry plasmids when reacted with appropriate att-containing PCR products formed linear recombinant molecules that could be easily visualized on an agarose gel when compared to control reactions containing no attL or attR PCR product. Thus, the functionality of Destination and Entry vectors constructed according to the invention may be determined either by carrying out the Destination or Entry recombination reactions as depicted in

-144-

Figures 2, 4, and 5A and 5B, or more rapidly by carrying out the linearization assay described in this Example.

Example 20: PCR Cloning Using Universal Adapter-Primers

As described herein, the cloning of PCR products using the GATEWAY™ PCR Cloning System (Life Technologies, Inc.; Rockville, MD) requires the addition of attB sites (attB1 and attB2) to the ends of gene-specific primers used in the PCR reaction. The protocols described in the preceding Examples suggest that the user add 29 bp (25 bp containing the attB site plus four G residues) to the gene-specific primer. It would be advantageous to high volume users of the GATEWAY™ PCR Cloning System to generate attB-containing PCR product using universal attB adapter-primers in combination with shorter gene-specific primers containing a specified overlap to the adapters. The following experiments demonstrate the utility of this strategy using universal attB adapter-primers and gene-specific primers containing overlaps of various lengths from 6 bp to 18 bp. The results demonstrate that gene-specific primers with overlaps of 10 bp to 18 bp can be used successfully in PCR amplifications with universal attB adapter-primers to generate full-length PCR products. These PCR products can then be successfully cloned with high fidelity in a specified orientation using the GATEWAY™ PCR Cloning System.

Methods and Results:

To demonstrate that universal attB adapter-primers can be used with gene-specific primers containing partial attB sites in PCR reactions to generate full-length PCR product, a small 256 bp region of the human hemoglobin cDNA was chosen as a target so that intermediate sized products could be distinguished from full-length products by agarose gel electrophoresis.

The following oligonucleotides were used:

B1-Hgb: GGGG ACA AGT TTG TAC AAA AAA GCA GGC T-5'-Hgb*
B2-Hgb: GGGG ACC ACT TTG TAC AAG AAA GCT GGG T-3'-Hgb**

-145-

	18B1-Hgb:	TG TAC AAA AAA GCA GGC T-5'-Hgb
	18B2-Hgb:	TG TAC AAG AAA GCT GGG T-3'-Hgb
	15B1-Hgb:	AC AAA AAA GCA GGC T-5'-Hgb
	15B2-Hgb:	AC AAG AAA GCT GGG T-3'-Hgb
5	12B1-Hgb:	AA AAA GCA GGC T-5'-Hgb
	12B2-Hgb:	AG AAA GCT GGG T-3'-Hgb
	11B1-Hgb:	A AAA GCA GGC T-5'-Hgb
	11B2-Hgb:	G AAA GCT GGG T-3'-Hgb
	10B1-Hgb:	AAA GCA GGC T-5'-Hgb
10	10B2-Hgb:	AAA GCT GGG T-3'-Hgb
	9B1-Hgb:	AA GCA GGC T-5'-Hgb
	9B2-Hgb:	AA GCT GGG T-3'-Hgb
	8B1-Hgb:	A GCA GGC T-5'-Hgb
	8B2-Hgb:	A GCT GGG T-3'-Hgb
15	7B1-Hgb:	GCA GGC T-5'-Hgb
	7B2-Hgb:	GCT GGG T-3'-Hgb
	6B1-Hgb:	CA GGC T-5'-Hgb
	6B2-Hgb:	CT GGG T-3'-Hgb

20 attB1 adapter: GGGG ACA AGT TTG TAC AAA AAA GCA GGC T
attB2 adapter: GGGG ACC ACT TTG TAC AAG AAA GCT GGG T

* -5'-Hgb = GTC ACT AGC CTG TGG AGC AAG A

** -3'-Hgb = AGG ATG GCA GAG GGA GAC GAC A

25

The aim of these experiments was to develop a simple and efficient universal adapter PCR method to generate attB containing PCR products suitable for use in the GATEWAY™ PCR Cloning System. The reaction mixtures and thermocycling conditions should be simple and efficient so that the universal adapter PCR method could be routinely applicable to any PCR product cloning application.

30

PCR reaction conditions were initially found that could successfully amplify predominately full-length PCR product using gene-specific primers containing 18bp and 15 bp overlap with universal attB primers. These conditions are outlined below:

35

-146-

10 pmoles of gene-specific primers
10 pmoles of universal attB adapter-primers
1 ng of plasmid containing the human hemoglobin cDNA.
100 ng of human leukocyte cDNA library DNA.
5 μ l of 10x PLATINUM Taq HiFi® reaction buffer (Life Technologies, Inc.)
2 μ l of 50 mM MgSO₄
1 μ l of 10 mM dNTPs
0.2 μ l of PLATINUM Taq HiFi® (1.0 unit)
H₂O to 50 μ l total reaction volume

10

Cycling conditions:

15

25 x |
 95°C/5 min
 94°C/15 sec
 50°C/30 sec
 68°C/1 min
 68°C/5 min
 5°C/hold

20 To assess the efficiency of the method, 2 μ l (1/25) of the 50 μ l PCR reaction was electrophoresed in a 3 % Agarose-1000 gel. With overlaps of 12 bp or less, smaller intermediate products containing one or no universal attB adapter predominated the reactions. Further optimization of PCR reaction conditions was obtained by titrating the amounts of gene-specific primers and universal attB adapter-primers. The PCR reactions were set up as outlined above except that the
25 amounts of primers added were:

0, 1, 3 or 10 pmoles of gene-specific primers
0, 10, 30 or 100 pmoles of adapter-primers

Cycling conditions:

5 25 x | 95°C/3 min
 | 94°C/15 sec
 | 50°C/45 sec
 | 68°C/1 min
 | 68°C/5 min
 | 5°C/hold

10 The use of limiting amounts of gene-specific primers (3 pmoles) and excess adapter-primers (30 pmoles) reduced the amounts of smaller intermediate products. Using these reaction conditions the overlap necessary to obtain predominately full-length PCR product was reduced to 12 bp. The amounts of gene-specific and adapter-primers was further optimized in the following PCR reactions:

15

0, 1, 2 or 3 pmoles of gene-specific primers
0, 30, 40 or 50 pmoles of adapter-primers

Cycling conditions:

20 25 x | 95°C/3 min
 | 94°C/15 sec
 | 48°C/1 min
 | 68°C/1 min
 | 68°C/5 min
25 | 5°C/hold

30

The use of 2 pmoles of gene-specific primers and 40 pmoles of adapter-primers further reduced the amounts of intermediate products and generated predominately full-length PCR products with gene-specific primers containing an 11 bp overlap. The success of the PCR reactions can be assessed in any PCR application by performing a no adapter control. The use of limiting amounts of gene-specific primers should give faint or barely visible bands when 1/25 to 1/10 of the PCR reaction is electrophoresed on a standard agarose gel. Addition of the

universal attB adapter-primers should generate a robust PCR reaction with a much higher overall yield of product.

PCR products from reactions using the 18 bp, 15 bp, 12 bp, 11 bp and 10 bp overlap gene-specific primers were purified using the CONCERT® Rapid PCR Purification System (PCR products greater than 500 bp can be PEG precipitated). The purified PCR products were subsequently cloned into an attP containing plasmid vector using the GATEWAY™ PCR Cloning System (Life Technologies, Inc.; Rockville, MD) and transformed into *E. coli*. Colonies were selected and counted on the appropriate antibiotic media and screened by PCR for correct inserts and orientation.

Raw PCR products (unpurified) from the attB adapter PCR of a plasmid clone of part of the human beta-globin (Hgb) gene were also used in GATEWAY™ PCR Cloning System reactions. PCR products generated with the full attB B1/B2-Hgb, the 12B1/B2, 11B1/B2 and 10B1/B2 attB overlap Hgb primers were successfully cloned into the GATEWAY™ pENTR21 attP vector (Figure 49). 24 colonies from each ($24 \times 4 = 96$ total) were tested and each was verified by PCR to contain correct inserts. The cloning efficiency expressed as cfu/ml is shown below:

Primer Used	cfu/ml
Hgb full attB	8,700
Hgb 12 bp overlap	21,000
Hgb 11 bp overlap	20,500
Hgb 10 bp overlap	13,500
GFP control	1,300

Interestingly, the overlap PCR products cloned with higher efficiency than did the full attB PCR product. Presumably, and as verified by visualization on agarose gel, the adapter PCR products were slightly cleaner than was the full attB PCR product. The differences in colony output may also reflect the proportion of PCR product molecules with intact attB sites.

Using the attB adapter PCR method, PCR primers with 12 bp attB overlaps were used to amplify cDNAs of different sizes (ranging from 1 to 4 kb)

from a leukocyte cDNA library and from first strand cDNA prepared from HeLa total RNA. While three of the four cDNAs were able to be amplified by this method, a non-specific amplification product was also observed that under some conditions would interfere with the gene-specific amplification. This non-specific product was amplified in reactions containing the attB adapter-primers alone without any gene-specific overlap primers present. The non-specific amplification product was reduced by increasing the stringency of the PCR reaction and lowering the attB adapter PCR primer concentration.

These results indicate that the adapter-primer PCR approach described in this Example will work well for cloned genes. These results also demonstrate the development of a simple and efficient method to amplify PCR products that are compatible with the GATEWAY™ PCR Cloning System that allows the use of shorter gene-specific primers that partially overlap universal attB adapter-primers. In routine PCR cloning applications, the use of 12 bp overlaps is recommended. The methods described in this Example can thus reduce the length of gene-specific primers by up to 17 residues or more, resulting in a significant savings in oligonucleotide costs for high volume users of the GATEWAY™ PCR Cloning System. In addition, using the methods and assays described in this Example, one of ordinary skill can, using only routine experimentation, design and use analogous primer-adapters based on or containing other recombination sites or fragments thereof, such as *attL*, *attR*, *attP*, *lox*, FRT, etc.

Example 21: Mutational Analysis of the Bacteriophage Lambda *attL* and *attR* Sites: Determinants of *att* Site Specificity in Site-specific Recombination

To investigate the determinants of *att* site specificity, the bacteriophage lambda *attL* and *attR* sites were systematically mutagenized. As noted herein, the determinants of specificity have previously been localized to the 7 bp overlap region (TTTATAC, which is defined by the cut sites for the integrase protein and is the region where strand exchange takes place) within the 15 bp core region (GCTTTTTTATACTAA) which is identical in all four lambda *att* sites, *attB*, *attP*, *attL* and *attR*. This core region, however, has not heretofore been systematically

-150-

mutagenized and examined to define precisely which mutations produce unique changes in *att* site specificity.

Therefore, to examine the effect of *att* sequence on site specificity, mutant *attL* and *attR* sites were generated by PCR and tested in an *in vitro* site-specific recombination assay. In this way all possible single base pair changes within the 7 bp overlap region of the core *att* site were generated as well as five additional changes outside the 7 bp overlap but within the 15 bp core *att* site. Each *attL* PCR substrate was tested in the *in vitro* recombination assay with each of the *attR* PCR substrates.

10

Methods

15

To examine both the efficiency and specificity of recombination of mutant *attL* and *attR* sites, a simple *in vitro* site-specific recombination assay was developed. Since the core regions of *attL* and *attR* lie near the ends of these sites, it was possible to incorporate the desired nucleotide base changes within PCR primers and generate a series of PCR products containing mutant *attL* and *attR* sites. PCR products containing *attL* and *attR* sites were used as substrates in an *in vitro* reaction with GATEWAY™ LR Clonase™ Enzyme Mix (Life Technologies, Inc.; Rockville, MD). Recombination between a 1.5 kb *attL* PCR product and a 1.0 kb *attR* PCR product resulted in a 2.5 kb recombinant molecule that was monitored using agarose gel electrophoresis and ethidium bromide staining.

20

Plasmid templates pEZC1301 (Figure 84) and pEZC1313 (Figure 85), each containing a single wild type *attL* or *attR* site, respectively, were used for the generation of recombination substrates. The following list shows primers that were used in PCR reactions to generate the *attL* PCR products that were used as substrates in L x R Clonase reactions (capital letters represent changes from the wild-type sequence, and the underline represents the 7 bp overlap region within the 15 bp core *att* site; a similar set of PCR primers was used to prepare the *attR* PCR products containing matching mutations):

-151-

GATEWAY™ sites (note: attL2 sequence in GATEWAY™ plasmids begins "acccca" while the attL2 site in this example begins "agcct" to reflect wild-type attL outside the core region.):

5

attL1: gggg agcct gctttttGtacAaa gttggcatta taaaaa-
agca ttgc

10

attL2: gggg agcct gctttCttGtacAaa gttggcatta taaaaa-
agca ttgc

Wild-type:

attL0: gggg agcct gctttttataactaa gttggcatta taaaaa-
agca ttgc

15

Single base changes from wild-type:

attLT1A: gggg agcct gctttAttataactaa gttggcatta taaaaa-
agca ttgc

20

attLT1C: gggg agcct gctttCttataactaa gttggcatta taaaaa-
agca ttgc

attLT1G: gggg agcct gctttGttataactaa gttggcatta taaaaa-
agca ttgc

25

attLT2A: gggg agcct gctttAtataactaa gttggcatta taaaaa-
agca ttgc

30

attLT2C: gggg agcct gctttCtataactaa gttggcatta taaaaa-
agca ttgc

attLT2G: gggg agcct gctttGtataactaa gttggcatta taaaaa-
aagca ttgc

35

-152-

attLT3A: gggg agcct gcttttAataactaa gttggcatta taaaa-
aagca ttgc

5 attLT3C: gggg agcct gcttttCataactaa gttggcatta taaaa-
aagca ttgc

10 attLT3G: gggg agcct gcttttGataactaa gttggcatta taaaa-
aagca ttgc

15 attLA4C: gggg agcct gctttttCtactaa gttggcatta taaaa-
aagca ttgc

20 attLA4T: gggg agcct gctttttTtactaa gttggcatta taaaa-
aagca ttgc

25 attLT5A: gggg agcct gcttttttaAactaa gttggcatta taaaa-
aagca ttgc

30 attLT5C: gggg agcct gcttttttaCactaa gttggcatta taaaa-
aagca ttgc

35 attLT5G: gggg agcct gcttttttaGactaa gttggcatta taaaa-
aagca ttgc

attLA6C: gggg agcct gctttttatCctaa gttggcatta taaaa-
aagca ttgc

-153-

attLA6G: gggg agcct gctttttatGctaa gttggcatta taaaa-
aagca ttgc

5 attLA6T: gggg agcct gctttttatTctaa gttggcatta taaaa-
aagca ttgc

10 attLC7A: gggg agcct gctttttataAataa gttggcatta taaaa-
aagca ttgc

15 attLC7G: gggg agcct gctttttataGtaa gttggcatta taaaa-
aagca ttgc

attLC7T: gggg agcct gctttttataTtaa gttggcatta taaaa-
aagca ttgc

Single base changes outside of the 7 bp overlap:

20 attL8: gggg agcct Actttttataactaa gttggcatta taaaa-
aagca ttgc

25 attL9: gggg agcct gcCttttataactaa gttggcatta taaaaaa-
agca ttgc

attL10: gggg agcct gcttCtttataactaa gttggcatta taaaaaa-
agca ttgc

30 attL14: gggg agcct gctttttatacCaa gttggcatta taaaaaa-
agca ttgc

35 attL15: gggg agcct gctttttataactaG gttggcatta taaaaaa-
agca ttgc

-154-

Note: additional vectors wherein the first nine bases are gggg agcca (i.e., substituting an adenine for the thymine in the position immediately preceding the 15-bp core region), which may or may not contain the single base pair substitutions (or deletions) outlined above, can also be used in these experiments.

5

Recombination reactions of *attL*- and *attR*-containing PCR products was performed as follows:

- 10 8 µl of H₂O
 2 µl of *attL* PCR product (100 ng)
 2 µl of *attR* PCR product (100 ng)
 4 µl of 5x buffer
 4 µl of GATEWAY™ LR Clonase™ Enzyme Mix
 20 µl total volume

15

Clonase reactions were incubated at 25°C for 2 hours.

20 2 µl of 10X Clonase stop solution (proteinase K, 2 mg/ml) were added to stop the reaction.

25 10 µl were run on a 1 % agarose gel.

Results

Each *attL* PCR substrate was tested in the *in vitro* recombination assay with each of the *attR* PCR substrates. Changes within the first three positions of the 7 bp overlap (TTTATAC) strongly altered the specificity of recombination. These mutant *att* sites each recombined as well as the wild-type, but only with their cognate partner mutant; they did not recombine detectably with any other *att* site mutant. In contrast, changes in the last four positions (TTTATAC) only partially altered specificity; these mutants recombined with their cognate mutant as well as wild-type *att* sites and recombined partially with all other mutant *att* sites except for those having mutations in the first three positions of the 7 bp

-155-

overlap. Changes outside of the 7 bp overlap were found not to affect specificity of recombination, but some did influence the efficiency of recombination.

Based on these results, the following rules for *att* site specificity were determined:

- 5
 - Only changes within the 7 bp overlap affect specificity.
 - Changes within the first 3 positions strongly affect specificity.
 - Changes within the last 4 positions weakly affect specificity.

10 Mutations that affected the overall efficiency of the recombination reaction were also assessed by this method. In these experiments, a slightly increased (less than 2-fold) recombination efficiency with *att*L T1A and *att*LC7T substrates was observed when these substrates were reacted with their cognate *att*R partners. Also observed were mutations that decreased recombination efficiency (approximately 2-3 fold), including *att*LA6G, *att*L14 and *att*L15. These mutations presumably reflect changes that affect Int protein binding at the core *att* site.

15 The results of these experiments demonstrate that changes within the first three positions of the 7 bp overlap (TTTATAC) strongly altered the specificity of recombination (i.e., *att* sequences with one or more mutations in the first three thymidines would only recombine with their cognate partners and would not cross-react with any other *att* site mutation). In contrast, mutations in the last four positions (TTTATAC) only partially altered specificity (i.e., *att* sequences with one or more mutations in the last four base positions would cross-react partially with the wild-type *att* site and all other mutant *att* sites, except for those having mutations in one or more of the first three positions of the 7 bp overlap). Mutations outside of the 7 bp overlap were not found to affect specificity of recombination, but some were found to influence (i.e., to cause a decrease in) the efficiency of recombination.

25 ***Example 22: Discovery of Att Site Mutations That Increase the Cloning Efficiency of GATEWAY™ Cloning Reactions***

30 In experiments designed to understand the determinants of *att* site specificity, point mutations in the core region of *att*L were made. Nucleic acid molecules containing these mutated *att*L sequences were then reacted in an LR

reaction with nucleic acid molecules containing the cognate *attR* site (*i.e.*, an *attR* site containing a mutation corresponding to that in the *attL* site), and recombinational efficiency was determined as described above. Several mutations located in the core region of the *att* site were noted that either slightly increased (less than 2-fold) or decreased (between 2-4-fold) the efficiency of the recombination reaction (Table 3).

Table 3. Effects of attL mutations on Recombination Reactions.

	<u>Site</u>	<u>Sequence</u>	<u>Effect on Recombination</u>
	attL0	agcctgcttttataactaagggtggcatta	
	attL5	agcctgctttAttataactaagggtggcatta	slightly increased
	attL6	agcctgcttttataTtaagggtggcatta	slightly increased
15	attL13	agcctgcttttatGctaagggtggcatta	decreased
	attL14	agcctgcttttatacCaagggtggcatta	decreased
	attL15	agcctgcttttataactaGgtggcatta	decreased
20	consensus	CAACTTnnTnnnAnnAAGTTG	

It was also noted that these mutations presumably reflected changes that either increased or decreased, respectively, the relative affinity of the integrase protein for binding the core *att* site. A consensus sequence for an integrase core-binding site (CAACTTNNT) has been inferred in the literature but not directly tested (see, *e.g.*, Ross and Landy, *Cell* 33:261-272 (1983)). This consensus core integrase-binding sequence was established by comparing the sequences of each of the four core *att* sites found in *attP* and *attB* as well as the sequences of five non-*att* sites that resemble the core sequence and to which integrase has been shown to bind *in vitro*. These experiments suggest that many more *att* site mutations might be identified which increase the binding of integrase to the core *att* site and thus increase the efficiency of GATEWAY™ cloning reactions.

Example 23: Effects of Core Region Mutations on Recombination Efficiency

To directly compare the cloning efficiency of mutations in the att site core region, single base changes were made in the attB2 site of an attB1-TET-attB2 PCR product. Nucleic acid molecules containing these mutated attB2 sequences were then reacted in a BP reaction with nucleic acid molecules containing non-cognate attP sites (*i.e.*, wildtype attP2), and recombinational efficiency was determined as described above. The cloning efficiency of these mutant attB2 containing PCR products compared to standard attB1-TET-attB2 PCR product are shown in Table 4.

10

Table 4. Efficiency of Recombination With Mutated attB2 Sites.

	<u>Site</u>	<u>Sequence</u>	<u>Mutation</u>	<u>Cloning Efficiency</u>
15	attB0	tcaagttagtataaaaaaggcaggct		
	attB1	ggggacaagttgtacaaaaaggcaggct		
	attB2	ggggaccactttgtacaagaagctgggt		100%
20	attB2.1	ggggaAactttgtacaagaagctgggt	C→A	40%
	attB2.2	ggggacAAactttgtacaagaagctgggt	C→A	131%
	attB2.3	ggggaccCctttgtacaagaagctgggt	A→C	4%
	attB2.4	ggggaccaAtttgtacaagaagctgggt	C→A	11%
	attB2.5	ggggaccacGttgtacaagaagctgggt	T→G	4%
	attB2.6	ggggaccactGtgtacaagaagctgggt	T→G	6%
25	attB2.7	ggggaccacttGtgtacaagaagctgggt	T→G	1%
	attB2.8	ggggaccactttTtacaagaagctgggt	G→T	0.5%

As noted above, a single base change in the attB2.2 site increased the cloning efficiency of the attB1-TET-attB2.2 PCR product to 131% compared to the attB1-TET-attB2 PCR product. Interestingly, this mutation changes the integrase core binding site of attB2 to a sequence that matches more closely the proposed consensus sequence.

-158-

Additional experiments were performed to directly compare the cloning efficiency of an attB1-TET-attB2 PCR product with a PCR product that contained attB sites containing the proposed consensus sequence (*see Example 22*) of an integrase core binding site. The following attB sites were used to amplify attB-TET PCR products:

attB1 ggggacaagttgtacaaaaaaaggct
 attB1.6 ggggacaaCtttgtacaaaaaaagTTggct
 attB2 ggggaccactttgtacaqaaagctgggt
 attB2.10 ggggacAactttgtacaqaaagTtgggt

BP reactions were carried out between 300 ng (100 fmoles) of pDONR201 (Figure 49A) with 80 ng (80 fmoles) of attB-TET PCR product in a 20 μ l volume with incubation for 1.5 hrs at 25°C, creating pENTR201-TET Entry clones. A comparison of the cloning efficiencies of the above-noted attB sites in BP reactions is shown in Table 5.

Table 5. Cloning efficiency of BP Reactions.

PCR product	CFU/ml	Fold Increase
B1-tet-B2	7,500	
B1.6-tet-B2	12,000	1.6 x
B1-tet-B2.10	20,900	2.8 x
B1.6-tet-B2.10	30,100	4.0 x

These results demonstrate that attB PCR products containing sequences that perfectly match the proposed consensus sequence for integrase core binding sites can produce Entry clones with four-fold higher efficiency than standard Gateway attB1 and attB2 PCR products.

The entry clones produced above were then transferred to pDEST20 (Figure 40A) via LR reactions (300 ng (64 fmoles) pDEST20 mixed with 50 ng (77 fmoles) of the respective pENTR201-TET Entry clone in 20 μ l volume; incubated for 1 hr incubation at 25°C). The efficiencies of cloning for these reactions are compared in Table 6.

-159-

Table 6. Cloning Efficiency of LR Reactions.

pENTR201-TET x pDEST20	CFU/ml	Fold Increase
L1-tet-L2	5,800	
L1.6-tet-L2	8,000	1.4
L1-tet-L2.10	10,000	1.7
L1.6-tet-L2.10	9,300	1.6

These results demonstrate that the mutations introduced into attB1.6 and attB2.10 that transfer with the gene into entry clones slightly increase the efficiency of LR reactions. Thus, the present invention encompasses not only mutations in *attB* sites that increase recombination efficiency, but also to the corresponding mutations that result in the *attL* sites created by the BP reaction.

To examine the increased cloning efficiency of the attB1.6-TET-attB2.10 PCR product over a range of PCR product amounts, experiments analogous to those described above were performed in which the amount of attB PCR product was titrated into the reaction mixture. The results are shown in Table 7.

Table 7. Titration of attB PCR products.

Amount of attB PCR product (ng)	PCR product	CFU/ml	Fold Increase
20	attB1-TET-attB2	3,500	6.1
	attB1.6-TET-attB2.10	21,500	
50	attB1-TET-attB2	9,800	5.0
	attB1.6-TET-attB2.10	49,000	
100	attB1-TET-attB2	18,800	2.8
	attB1.6-TET-attB2.10	53,000	
200	attB1-TET-attB2	19,000	2.5
	attB1.6-TET-attB2.10	48,000	

These results demonstrate that as much as a six-fold increase in cloning efficiency is achieved with the attB1.6-TET-attB2.10 PCR product as compared to the standard attB1-TET-attB2 PCR product at the 20 ng amount.

Example 24: Determination of attB Sequence Requirements for Optimum Recombination Efficiency

To examine the sequence requirements for attB and to determine which attB sites would clone with the highest efficiency from populations of degenerate attB sites, a series of experiments was performed. Degenerate PCR primers were designed which contained five bases of degeneracy in the B-arm of the attB site. These degenerate sequences would thus transfer with the gene into Entry clone in BP reactions and subsequently be transferred with the gene into expression clones in LR reactions. The populations of degenerate attB and attL sites could thus be cycled from attB to attL back and forth for any number of cycles. By altering the reaction conditions at each transfer step (for example by decreasing the reaction time and/or decreasing the concentration of DNA) the reaction can be made increasingly more stringent at each cycle and thus enrich for populations of attB and attL sites that react more efficiently.

The following degenerate PCR primers were used to amplify a 500 bp fragment from pUC18 which contained the lacZ alpha fragment (only the attB portion of each primer is shown):

attB1 GGGG ACAAGTTGTACAAA AAAGC AGGCT
attB1n16-20 GGGG ACAAGTTGTACAAA nnnnn AGGCT
attB1n21-25 GGGG ACAAGTTGTACAAA AAAGC nnnnn

attB2 GGGG ACCACTTGTACAAG AAAGC TGGGT
attB2n16-20 GGGG ACCACTTGTACAAG nnnnn TGGGT
attB2n21-25 GGGG ACCACTTGTACAAG AAAGC nnnnn

The starting population size of degenerate att sites is 4^5 or 1024 molecules. Four different populations were transferred through two BP reactions and two LR reactions. Following transformation of each reaction, the population of transformants was amplified by growth in liquid media containing the appropriate selection antibiotic. DNA was prepared from the population of clones by alkaline

-161-

ysis miniprep and used in the next reaction. The results of the BP and LR cloning reactions are shown below.

BP-1, overnight reactions

	cfu/ml	percent of control
attB1-LacZa-attB2	78,500	100 %
attB1n16-20-LacZa-attB2	1,140	1.5 %
attB1n21-25-LacZa-attB2	11,100	14 %
attB1-LacZa-attB2n16-20	710	0.9 %
attB1-LacZa-attB2n21-25	16,600	21 %

LR-1, pENTR201-LacZa x pDEST20/EcoRI, 1hr reactions

	cfu/ml	percent of control
attL1-LacZa-attL2	20,000	100 %
attL1n16-20-LacZa-attL2	2,125	11 %
attL1n21-25-LacZa-attL2	2,920	15 %
attL1-LacZa-attL2n16-20	3,190	16 %
attL1-LacZa-attL2n21-25	1,405	7 %

BP-2, pEXP20-LacZa/Scal x pDONR 201, 1hr reactions

	cfu/ml	percent of control
attB1-LacZa-attB2	48,600	100 %
attB1n16-20-LacZa-attB2	22,800	47 %
attB1n21-25-LacZa-attB2	31,500	65 %
attB1-LacZa-attB2n16-20	42,400	87 %
attB1-LacZa-attB2n21-25	34,500	71 %

LR-2, pENTR201-LacZa x pDEST6/NcoI, 1hr reactions

	cfu/ml	percent of control
attL1-LacZa-attL2	23,000	100 %
attL1n16-20-LacZa-attL2	49,000	213 %
attL1n21-25-LacZa-attL2	18,000	80 %
attL1-LacZa-attL2n16-20	37,000	160 %
attL1-LacZa-attL2n21-25	57,000	250 %

These results demonstrate that at each successive transfer, the cloning efficiency of the entire population of att sites increases, and that there is a great deal of flexibility in the definition of an *attB* site. Specific clones may be isolated from the above reactions, tested individually for recombination efficiency, and

-162-

sequenced. Such new specificities may then be compared to known examples to guide the design of new sequences with new recombination specificities. In addition, based on the enrichment and screening protocols described herein, one of ordinary skill can easily identify and use sequences in other recombination sites, e.g., other *att* sites, *lox*, FRT, etc., that result in increased specificity in the recombination reactions using nucleic acid molecules containing such sequences.

5

Example 25: Design of att Site PCR Adapter-Primers

10 Additional studies were performed to design gene-specific primers with 12bp of attB1 and attB2 at their 5'-ends. The optimal primer design for *att*-containing primers is the same as for any PCR primers: the gene-specific portion of the primers should ideally have a Tm of > 50°C at 50 mM salt (calculation of Tm is based on the formula $59.9 + 41(\%GC) - 675/n$).

15

Primers:

12bp attB1: AA AAA GCA GGC TNN - forward gene-specific primer

20

12bp attB2: A GAA AGC TGG GTN - reverse gene-specific primer

attB1 adapter primer: GGGGACAAGTTGTACAAAAAAGCAGGCT

attB2 adapter primer: GGGGACCACTTGTACAAGAAAGCTGGGT

25

Protocol:

(1) Mix 200 ng of cDNA library or 1 ng of plasmid clone DNA (alternatively, genomic DNA or RNA could be used) with 10 pmoles of gene specific primers in a 50 µl PCR reaction, using one or more polypeptides having DNA polymerase activity such as those described herein. (The addition of greater than 10 pmoles of gene-specific primers can decrease the yield of attB PCR product. In addition, if RNA is used, a standard reverse transcriptase-PCR (RT-

30

-163-

PCR) protocol should be followed; see, e.g., Gerard, G.F., et al., *FOCUS* 11:60 (1989); Myers, T.W., and Gelfand, D.H., *Biochem.* 30:7661 (1991); Freeman, W.N., et al., *BioTechniques* 20:782 (1996); and U.S. Application No. 09/064,057, filed April 22, 1998, the disclosures of all of which are incorporated herein by reference.)

5

1st PCR profile:

- (a) 95°C for 3 minutes
- (b) 10 cycles of:
 - (i) 94°C for 15 seconds
 - (ii) 50°C* for 30 seconds
 - (iii) 68°C for 1 minute/kb of target amplicon
- (c) 68°C for 5 minutes
- (d) 10°C hold

10

*The optimal annealing temperature is determined by the calculated Tm of the gene-specific part of the primer.

15

(2) Transfer 10 µl to a 40 µl PCR reaction mix containing 35 pmoles each of the attB1 and attB2 adapter primers.

20

2nd PCR profile:

- (a) 95°C for 1 minute
- (b) 5 cycles of:
 - (i) 94°C for 15 seconds
 - (ii) 45°C* for 30 seconds
 - (iii) 68°C for 1 minute/kb of target amplicon
- (c) 15-20 cycles** of:
 - (i) 94°C for 15 seconds
 - (ii) 55°C* for 30 seconds

25

30

-164-

- (iii) 68°C for 1 minute/kb of target amplicon
(d) 68°C for 5 minutes
(e) 10°C hold

5 *The optimal annealing temperature is determined by the calculated Tm of the gene-specific part of the primer.
**15 cycles is sufficient for low complexity targets.

Notes:

- 10 1. It is useful to perform a no-adapter primer control to assess the yield of attB PCR product produced.
2. Linearized template usually results in slightly greater yield of PCR product.

15 ***Example 26: One-Tube Recombinational Cloning Using the GATEWAY™ Cloning System***

20 To provide for easier and more rapid cloning using the GATEWAY™ cloning system, we have designed a protocol whereby the BP and LR reactions may be performed in a single tube (a "one-tube" protocol). The following is an example of such a one-tube protocol; in this example, an aliquot of the BP reaction is taken before adding the LR components, but the BP and LR reactions may be performed in a one-tube protocol without first taking the BP aliquot:

<u>Reaction Component</u>	<u>Volume</u>
attB DNA (100-200 ng/25 µl reaction)	1-12.5 µl
attP DNA (pDONR201) 150 ng/µl	2.5 µl
5X BP Reaction Buffer	5.0 µl
Tris-EDTA	(to 20 µl)
<u>BP Clonase</u>	<u>5.0 µl</u>
Total vol.	25 µl

-165-

After the above components were mixed in a single tube, the reaction mixtures were incubated for 4 hours at 25°C. A 5 µl aliquot of reaction mixture was removed, and 0.5 µl of 10X stop solution was added to this reaction mixture and incubated for 10 minutes at 37°C. Competent cells were then transformed with 1-2 µl of the BP reaction per 100 µl of cells; this transformation yielded colonies of Entry Clones for isolation of individual Entry Clones and for quantitation of the BP Reaction efficiency.

To the remaining 20 µl of BP reaction mixture, the following components of the LR reaction were added:

<u>Reaction Component</u>	<u>Final Concentration</u>	<u>Volume Added</u>
NaCl	0.75 M	1 µl
Destination Vector	150 ng/µl	3 µl
<u>LR Clonase</u>		<u>6 µl</u>
Total vol.		30 µl

After the above components were mixed in a single tube, the reaction mixtures were incubated for 2 hours at 25°C. 3 µl of 10X stop solution was added, and the mixture was incubated for 10 minutes at 37°C. Competent cells were then transformed with 1-2 µl of the reaction mixture per 100 µl of cells

Notes:

1. If desired, the Destination Vector can be added to the initial BP reaction.
2. The reactions can be scaled down by 2x, if desired.
3. Shorter incubation times for the BP and/or LR reactions can be used (scaled to the desired cloning efficiencies of the reaction), but a lower number of colonies will typically result.
4. To increase the number of colonies obtained by several fold, incubate the BP reaction for 6-20 hours and increase the LR reaction to 3 hours. Electroporation also works well with 1-2 ul of the PK-treated reaction mixture.

-166-

5. PCR products greater than about 5 kb may show significantly lower cloning efficiency in the BP reaction. In this case, we recommend using a one-tube reaction with longer incubation times (e.g., 6-18 hours) for both the BP and LR steps.

5

Example 27: Relaxation of Destination Vectors During the LR Reaction

To further optimize the LR Reaction, the composition of the LR Reaction buffer was modified from that described above and this modified buffer was used in a protocol to examine the impact of enzymatic relaxation of Destination Vectors during the LR Reaction.

10 LR Reactions were set up as usual (see, e.g., Example 6), except that 5X BP Reaction Buffer (see Example 5) was used for the LR Reaction. To accomplish Destination Vector relaxation during the LR Reaction, Topoisomerase I (Life Technologies, Inc., Rockville, MD; Catalogue No. 38042-016) was added to the reaction mixture at a final concentration of ~15U per μ g of total DNA in the reaction (for example, for reaction mixtures with a total of 400ng DNA in the 20 μ l LR Reaction, ~6units of Topoisomerase I was added).

15 Reaction mixtures were set up as follows:

20

<u>Reaction Component</u>	<u>Volume</u>
ddH ₂ O	6.5 μ l
4X BP Reaction Buffer	5 μ l
100ng single chain/linear pENTR CAT, 50 ng/ μ l	2 μ l
25 300ng single chain/linear pDEST6, 150ng/ μ l	2 μ l
Topoisomerase I, 15 U/ml	0.5 μ l
LR Clonase	4 μ l

25

30 Reaction mixtures were incubated at 25°C for 1hour, and 2 μ l of 2 μ g/ μ l Proteinase K was then added and mixtures incubated for 10 minutes at 37°C to stop the LR Reaction. Competent cells were then transformed as described in the preceding examples. The results of these studies demonstrated that relaxation of

substrates in the LR reaction using Topoisomerase I resulted in a 2- to 10-fold increase in colony output compared to those LR reactions performed without including Topoisomerase I.

Having now fully described the present invention in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious to one of ordinary skill in the art that the same can be performed by modifying or changing the invention within a wide and equivalent range of conditions, formulations and other parameters without affecting the scope of the invention or any specific embodiment thereof, and that such modifications or changes are intended to be encompassed within the scope of the appended claims.

All publications, patents and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains, and are herein incorporated by reference to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference.

167.1

Applicant's or agent's file reference number	0942.~8PC03	International application No. t ¹ PCT/US 00/05432
--	-------------	---

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL
(PCT Rule 13bis)**

REC'D 17 APR 2000

WIPO PCT

A. The indications made below relate to the microorganism referred to in the description on page 52, line 31.

B. IDENTIFICATION OF DEPOSITFurther deposits are identified on an additional sheet

Name of depositary institution

Agricultural Research Culture Collection (NRRL)
International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit

February 27, 1999

Accession Number

NRRL B-30099

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*)This information is continued on an additional sheet

Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

This sheet was received with the international application

Authorized officer

Signature: *[Signature]*
Name: M.P. Tacka I
Title: PCT Admin.
Phone: (708) 563-2230 (F2)

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

167.2

Applicant's or agent's file reference number	0942.468PC03	International application No. t ₁ PCT/US 00/05432
--	--------------	---

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)

REC'D 17 APR 2000

WIPO

PCT

- A. The indications made below relate to the microorganism referred to in the description on page 55, line 16.

B. IDENTIFICATION OF DEPOSITFurther deposits are identified on an additional sheet

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depository Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit
February 27, 1999

Accession Number
NRRL B-30100

C. ADDITIONAL INDICATIONS (leave blank if not applicable)This information is continued on an additional sheet

Escherichia coli DB3.1(pENTR-1A)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)**E. SEPARATE FURNISHING OF INDICATIONS** (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only

For International Bureau use only

This sheet was received with the international application

This sheet was received by the International Bureau on:

Authorized officer

Barbara Frisch
U.S. Patent and Trademark Office
1100 L Street, NW
Washington, DC 20591-0000

Authorized officer

167 3	International application No. tb- PCT/US00 00/05432
Applicant's or agent's file reference number	0942.468PC03

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL
(PCT Rule 13bis)**

REC'D : 11:30 AM 10/10/00

A. The indications made below relate to the microorganism referred to in the description on page 16.	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (<i>including postal code and country</i>) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30101
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>) This information is continued on an additional sheet <input type="checkbox"/> Escherichia coli DB3.1(pENTR-2B)	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)	
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the international Bureau later (<i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i>)	

For receiving Office use only		For International Bureau use only	
<input checked="" type="checkbox"/> This sheet was received with the international application 		<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	Marcela Flores IT Operations, USPTO Phoenix Telephone: (602) 273-3230 (F)		
Authorized officer			

167.4

Applicant's or agent's file reference number	0942.468PC03	International Application No. t _L PCT/US 0 / 05432
--	--------------	--

REC'D 17 APR 2000

WIPO PCT

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL
(PCT Rule 13bis)**

A. The indications made below relate to the microorganism referred to in the description on page 55, line 16.

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit February 27, 1999	Accession Number NRRL B-30102
--------------------------------------	----------------------------------

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*)

This information is continued on an additional sheet

Escherichia coli DB3.1(pENTR-3C)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

This sheet was received with the international application

Authorized officer

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

167.5

Applicant's or agent's file reference number	0942.468PC03	International application No. tbc PCT/US 00/05432
--	--------------	--

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL
(PCT Rule 13bis)**

REF. 17	17
V	PCT

A. The indications made below relate to the microorganism referred to in the description on page <u>8</u> .	REC'D 17 APR 2000
	WIPO PCT

B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	

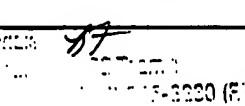
Address of depositary institution (<i>including postal code and country</i>) 1815 N. University Street Peoria, Illinois 61604 United States of America	
---	--

Date of deposit February 27, 1999	Accession Number NRRL B-30103
--------------------------------------	----------------------------------

C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)	This information is continued on an additional sheet <input type="checkbox"/>
Escherichia coli DB3.1(pEZC15101)	

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)
--

E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i>)

For receiving Office use only		For International Bureau use only	
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	 1202 PC03 1202 PC03 1202 PC03 1202 PC03 1202 PC03		
		Authorized officer	

167.6

Applicant's or agent's file reference number	0942.468PC03	International application No. 1. PCT/US 10/05432
--	--------------	---

REC'D 17

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL VPO**
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>54</u> , line <u>9</u> .	
B. IDENTIFICATION OF DEPOSIT	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Address of depositary institution (<i>including postal code and country</i>) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30104
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)	
This information is continued on an additional sheet <input type="checkbox"/> Escherichia coli DB3.1(pEZC15102)	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)	
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>)	
The indications listed below will be submitted to the international Bureau later (<i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i>)	

For receiving Office use only		For International Bureau use only	
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	<i>[Signature]</i>	Authorized officer	

167.7

Applicant's or agent's file reference number	0942.468PC03	International application No. tt PCT/US 00/05432
--	--------------	---

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)

DESM	17	ARR	100
V		T	

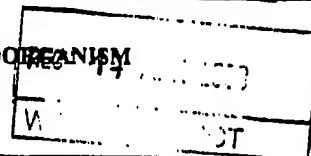
A. The indications made below relate to the microorganism referred to in the description on page <u>54</u> , line <u>9</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (<i>including postal code and country</i>) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30105
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>) This information is continued on an additional sheet <input type="checkbox"/> Escherichia coli DB3.1(pEYC15103)	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)	
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) The indications listed below will be submitted to the international Bureau later (<i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i>)	

For receiving Office use only		For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer	<i>[Signature]</i>	

167.8

Applicant's or agent's file reference number	0942.408PC03	International application No. tl PCT/US 00/05432
--	--------------	---

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)



A. The indications made below relate to the microorganism referred to in the description on page 51, line 20-21.

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit February 27, 1999	Accession Number NRRL B-30108
--------------------------------------	----------------------------------

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*)

This information is continued on an additional sheet

Escherichia coli DB10B(pCMV Sport6)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*If the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the international Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

For International Bureau use only

This sheet was received with the international application

This sheet was received by the International Bureau on:

Authorized officer

Barbara Frisch *BF*

Authorized officer

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group of nucleotide sequences consisting of an attB1 nucleotide sequence as set forth in Figure 9, an attB2 nucleotide sequence as set forth in Figure 9, an attP1 nucleotide sequence as set forth in Figure 9, an attP2 nucleotide sequence as set forth in Figure 9, an attL1 nucleotide sequence as set forth in Figure 9, an attL2 nucleotide sequence as set forth in Figure 9, an attR1 nucleotide sequence as set forth in Figure 9, an attR2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, and a mutant, fragment, or derivative thereof.
2. An isolated nucleic acid molecule comprising an attB1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.
3. An isolated nucleic acid molecule comprising an attB2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.
4. An isolated nucleic acid molecule comprising an attP1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.
5. An isolated nucleic acid molecule comprising an attP2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.
6. An isolated nucleic acid molecule comprising an attL1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

7. An isolated nucleic acid molecule comprising an attL2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

5 8. An isolated nucleic acid molecule comprising an attR1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

10 9. An isolated nucleic acid molecule comprising an attR2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

15 10. The isolated nucleic acid molecule of claim 1, further comprising one or more functional or structural nucleotide sequences selected from the group consisting of one or more multiple cloning sites, one or more localization signals, one or more transcription termination sites, one or more transcriptional regulatory sequences, one or more translational signals, one or more origins of replication, one or more fusion partner peptide-encoding nucleic acid molecules, one or more protease cleavage sites, and one or more 5' polynucleotide extensions.

20 11. The nucleic acid molecule of claim 10, wherein said transcriptional regulatory sequence is a promoter, an enhancer, or a repressor.

25 12. The nucleic acid molecule of claim 10, wherein said fusion partner peptide-encoding nucleic acid molecule encodes glutathione S-transferase (GST), hexahistidine (His₆), or thioredoxin (Trx).

13. The nucleic acid molecule of claim 10, wherein said 5' polynucleotide extension consists of from one to five nucleotide bases.

30 14. The nucleic acid molecule of claim 13, wherein said 5' polynucleotide extension consists of four or five guanine nucleotide bases.

-170-

15. A primer nucleic acid molecule suitable for amplifying a target nucleotide sequence, comprising the isolated nucleic acid molecule of claim 1 or a portion thereof linked to a target-specific nucleotide sequence useful in amplifying said target nucleotide sequence.

5

16. The primer nucleic acid molecule of claim 15, wherein said primer comprises an attB1 nucleotide sequence having the sequence shown in Figure 9 or a portion thereof, or a polynucleotide complementary to the sequence shown in Figure 9 or a portion thereof.

10

17. The primer nucleic acid molecule of claim 15, wherein said primer comprises an attB2 nucleotide sequence having the sequence shown in Figure 9 or a portion thereof, or a polynucleotide complementary to the sequence shown in Figure 9 or a portion thereof.

15

18. The primer nucleic acid molecule of claim 15, further comprising a 5' terminal extension of four or five guanine bases.

20. 19. A vector comprising the isolated nucleic acid molecule of claim 1.

20. 20. The vector of claim 19, wherein said vector is an Expression Vector.

25. 21. A host cell comprising the isolated nucleic acid molecule of claim 1 or the vector of claim 19.

22. 22. A method of synthesizing or amplifying one or more nucleic acid molecules comprising:

30. (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template-specific sequence that is complementary to or capable of hybridizing to said

templates and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said second primer is homologous to or complementary to at least a portion of said first primer; and

- 5 (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one or both termini of said molecules.

10

23. A method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- 15 (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template-specific sequence that is complementary to or capable of hybridizing to said templates and at least a portion of a recombination site, and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said recombination site on said second primer is complementary to or

20

homologous to at least a portion of said recombination site on said first primer; and

25

- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one or both termini of said molecules.

30

24. A method of amplifying or synthesizing one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity

and one or more first primers comprising at least a portion of a recombination site and a template-specific sequence that is complementary to or capable of hybridizing to said template;

- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more first nucleic acid molecules complementary to all or a portion of said templates wherein said molecules comprise at least a portion of a recombination site at one or both termini of said molecules;
 - (c) mixing said molecules with one or more second primers comprising one or more recombination sites, wherein said recombination sites of said second primers are homologous to or complementary to at least a portion of said recombination sites on said first nucleic acid molecules; and
 - (d) incubating said mixture under conditions sufficient to synthesize or amplify one or more second nucleic acid molecules complementary to all or a portion of said first nucleic acid molecules and which comprise one or more recombination sites at one or both termini of said molecules.

10

15

20

25. A polypeptide encoded by the isolated nucleic acid molecule of any one of claims 1-10.

25

26. An isolated nucleic acid molecule comprising one or more *att* recombination sites comprising at least one mutation in its core region that increases the specificity of interaction between said recombination site and a second *att* recombination site.

30

27. The isolated nucleic acid molecule of claim 26, wherein said mutation is at least one substitution mutation of at least one nucleotide in the seven basepair overlap region of said core region of said recombination site

5 28. The isolated nucleic acid molecule of claim 26, wherein said nucleic acid molecule comprises the sequence NNNATAC, wherein "N" refers to any nucleotide with the proviso that if one of the first three nucleotides in the consensus sequence is a T/U, then at least one of the other two of the first three nucleotides is not a T/U.

10 29. An isolated nucleic acid molecule comprising one or more mutated *att* recombination sites comprising at least one mutation in its core region that enhances the efficiency of recombination between a first nucleic acid molecule comprising said mutated *att* recombination site and a second nucleic acid molecule comprising a second recombination site that interacts with said mutated *att* recombination site.

15 30. The isolated nucleic acid molecule of claim 29, wherein said mutated *att* recombination site is a mutated *attL* site comprising a core region having the nucleotide sequence caactnnntnnnannaagttg, wherein "n" represents any nucleotide.

20 31. The isolated nucleic acid molecule of claim 30, wherein said mutated *attL* recombination site comprises a core region having a nucleotide sequence selected from agcctgcttattatactaagttggcatta (*attL5*) and agcctgcttttatattaagtggcatta (*attL6*).

25 32. The isolated nucleic acid molecule of claim 29, wherein said mutated *att* recombination site comprises a core region having a nucleotide sequence selected from the group consisting of ggggacaacttgtacaaaaaagttggct (*attB1.6*), ggggacaacttgtacaagaaagctgggt (*attB2.2*), and ggggacaacttgtacaagaaagttgggt (*attB2.10*).

30 33. A vector selected from the group consisting of pENTR1A, pENTR2B, pENTR3C, pENTR4, pENTR5, pENTR6, pENTR7, pENTR8, pENTR9, pENTR10, pENTR11, pDEST1, pDEST2, pDEST3, pDEST4,

-174-

pDEST5, pDEST6, pDEST7, pDEST8, pDEST9, pDEST10, pDEST11, pDEST12.2 (also known as pDEST12), pDEST13, pDEST14, pDEST15, pDEST16, pDEST17, pDEST18, pDEST19, pDEST20, pDEST21, pDEST22, pDEST23, pDEST24, pDEST25, pDEST26, pDEST27, pDEST28, pDEST29, pDEST30, pDEST31, pDEST32, pDEST33, pDEST34, pDONR201 (also known as pENTR21 attP vector or pAttPkan Donor Vector), pDONR202, pDONR203 (also known as pEZ15812), pDONR204, pDONR205, pDONR206 (also known as pENTR22 attP vector or pAttPgen Donor Vector), pDONR207, pMAB58, pMAB62, pMAB85 and pMAB86.

5

10

34. A host cell comprising the vector of claim 33.

35. A polypeptide encoded by the vector of claim 33.

15

36. A kit for use in synthesizing a nucleic acid molecule, said kit comprising the isolated nucleic acid molecule of any one of claims 1-10, 26 and 29.

20

37. A kit for use in synthesizing a nucleic acid molecule, said kit comprising the primer of claim 15 or claim 18.

38. A kit for use in cloning a nucleic acid molecule, said kit comprising the vector of claim 19 or claim 33.

1/240

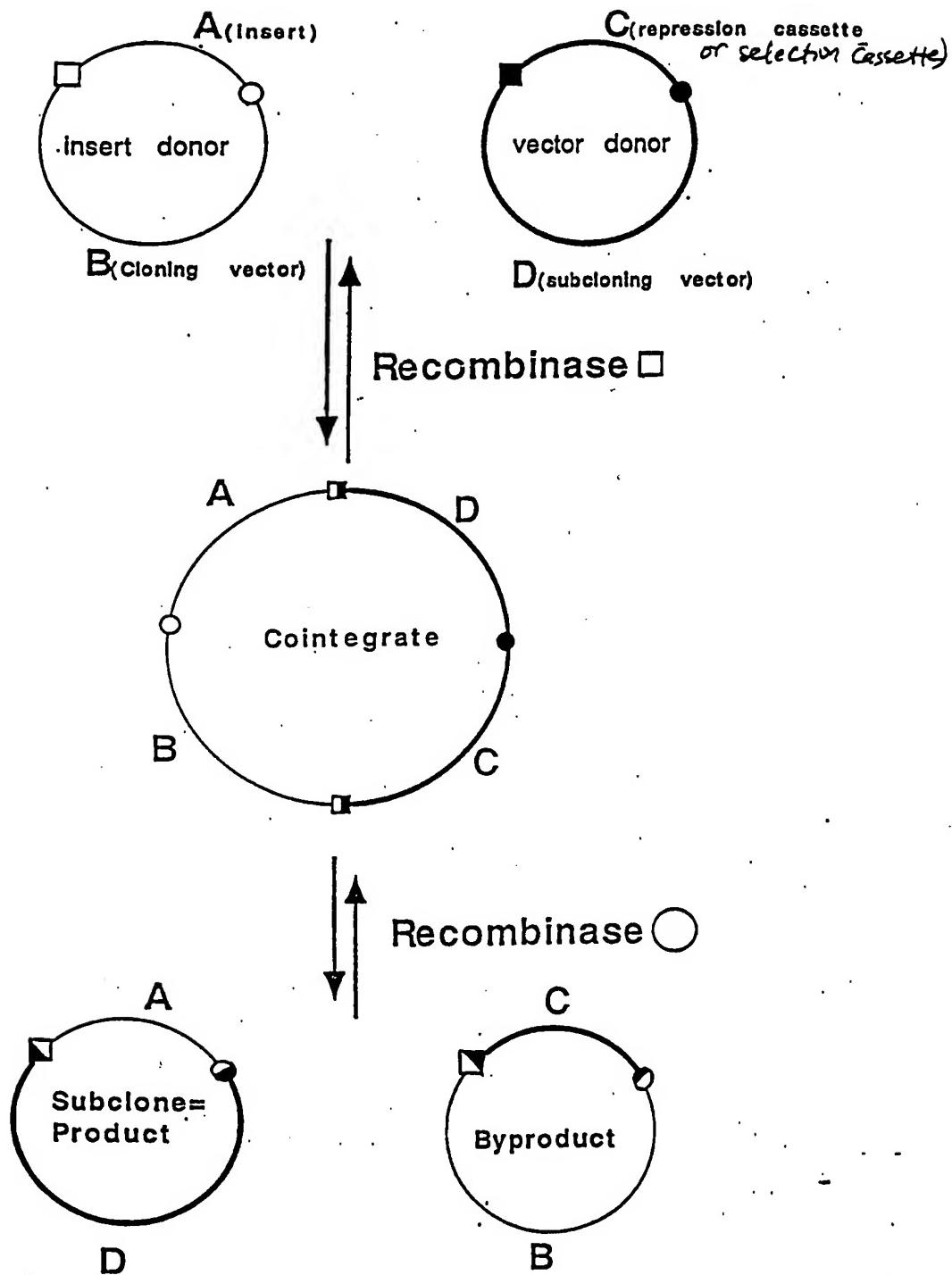


Figure 1

2/240

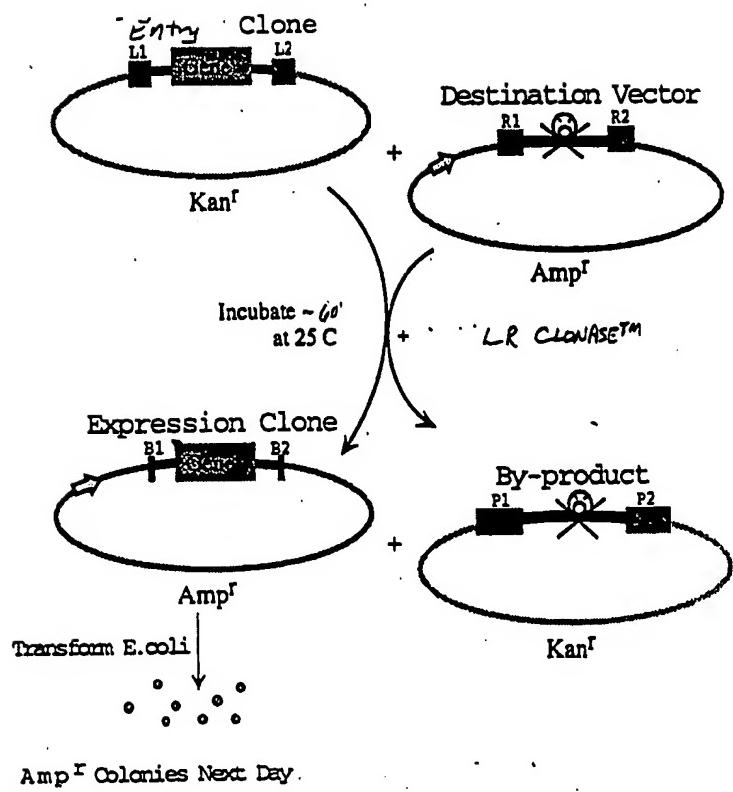


FIGURE 2

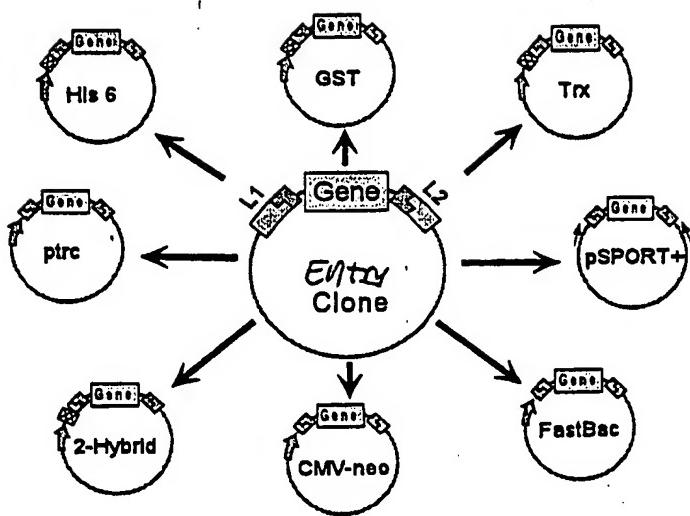


FIGURE 3

4/24/0

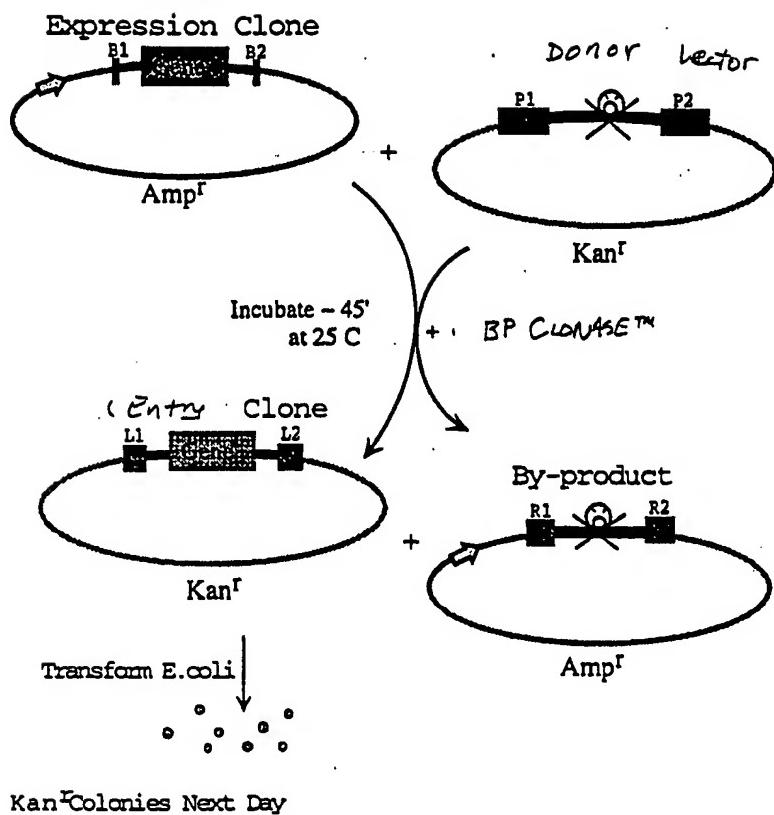


FIGURE 4

5/240

A

B

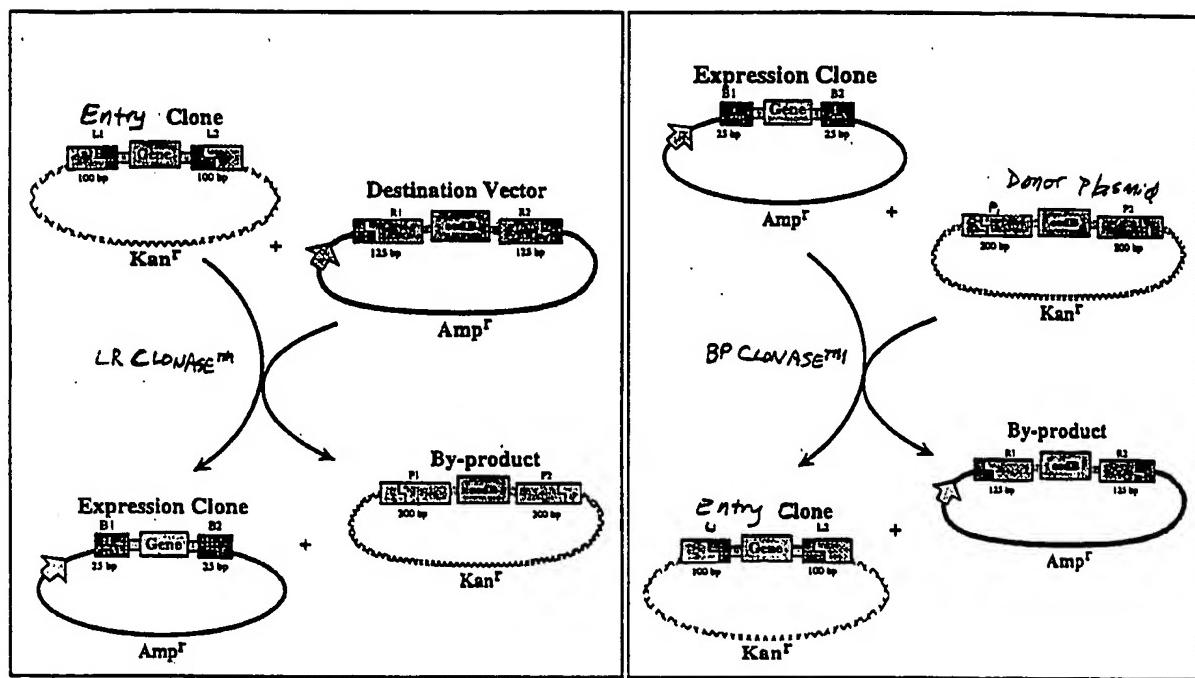


FIGURE 5

6/240

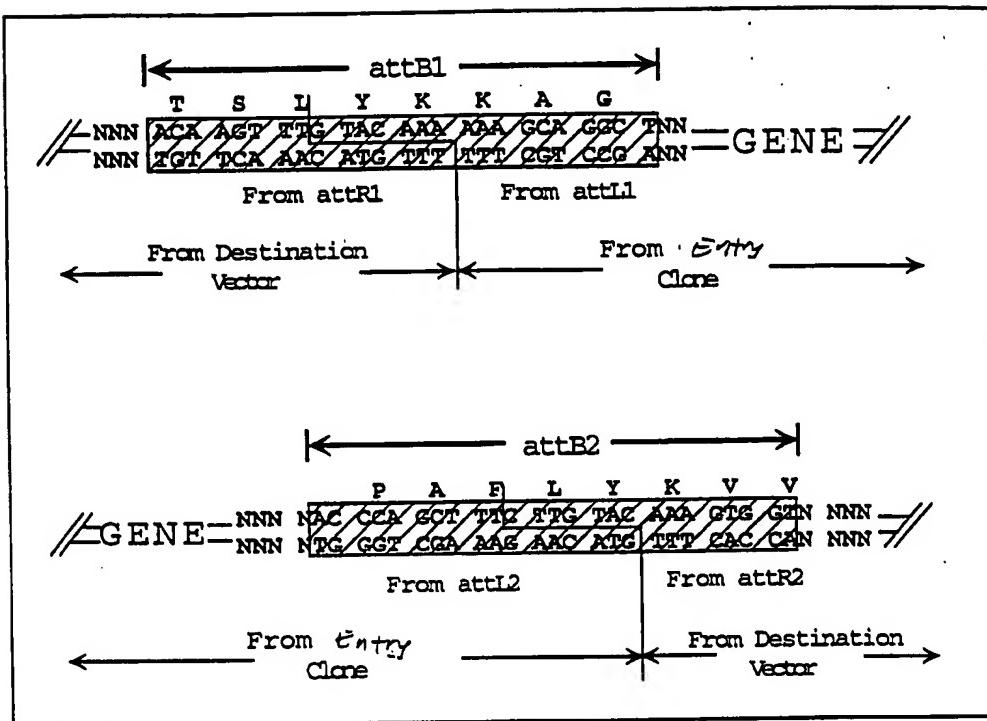


FIGURE 6

7/240

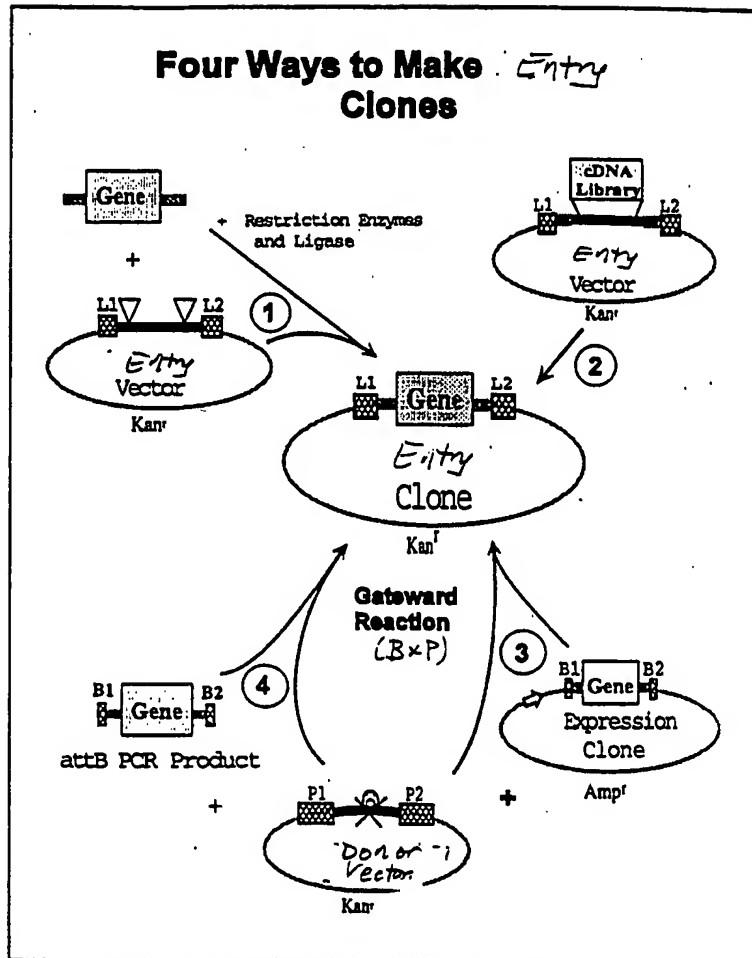


FIGURE 7

8/260

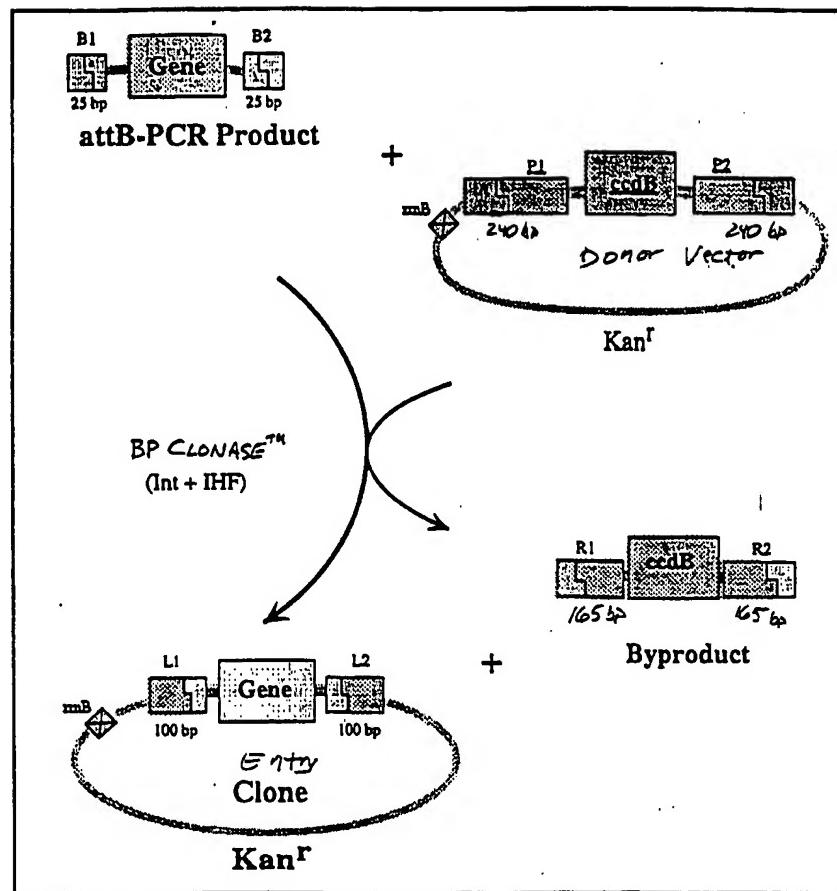


FIGURE 8

Recombination Site Nucleotide Sequences

attB1: 5'-ACAAGTTGTACAAAAAAGCAGGCT-3'

attB2: 5'-ACCCAGCTTCTTGTACAAAGTGGT-3'

attP1: 5'-TACAGGTCACTAACCATCTAAGTAGTTGATTGACTGGATATG-TTGTGTTTACAGTATTATGTAGTCTGTTTTATGCAAAATCTAATTAT-ATATATTGATATTTATATCATTACGTTCTCGTCAGCTTTTGAC-AAAGTGGCATTATAAAAAGCATTGCTCATCAATTGTTGCAACGAACA-GGTCACTATCAGTCAAAATAAAATCATTATTG-3'

attP2: 5'-CAAATAATGATTTATTTGACTGATAGTGACCTGTTGCAACAAAT-TGATAAGCAATGCTTCTTATAATGCCAACTTGACAGAAAGCTGAAC-GAGAAACGTAAAATGATATAAATATCAATATATTAAATTAGATTTGCAT-AAAAAACAGACTACATAACTGTAAAACACAACATATCCAGTCACTATGA-ATCAACTACTTAGATGGTATTAGTGACCTGTA-3'

attR1: 5'-ACAAGTTGTACAAAAAAGCTGAACGAGAAACGTAAAATGATATAAA-TATCAATATATAAATTAGATTTGCATAAAAACAGACTACATAATAC-TGTAAAACACAACATATCCAGTCACTATG-3'

attR2: 5'-GCAGGTGACCATAGTGACTGGATATGTTGTTTACAGTATTAT-GTAGTCTGTTTTATGCAAAATCTAATTAAATATATTGATATTT-ATATCATTACGTTCTCGTCAGCTTCTGTACAAAGTGGT-3'

attL1: 5'-CAAATAATGATTTATTTGACTGATAGTGACCTGTTGCAAC-AAAATTGATAAGCAATGCTTTTATAATGCCAACTTGACAAAAAA-GCAGGCT-3'

attL2: 5'-CAAATAATGATTTATTTGACTGATAGTGACCTGTTGCAACAA-ATTGATAAGCAATGCTTCTTATAATGCCAACTTGACAAAGAAAGCTGGGT-3'

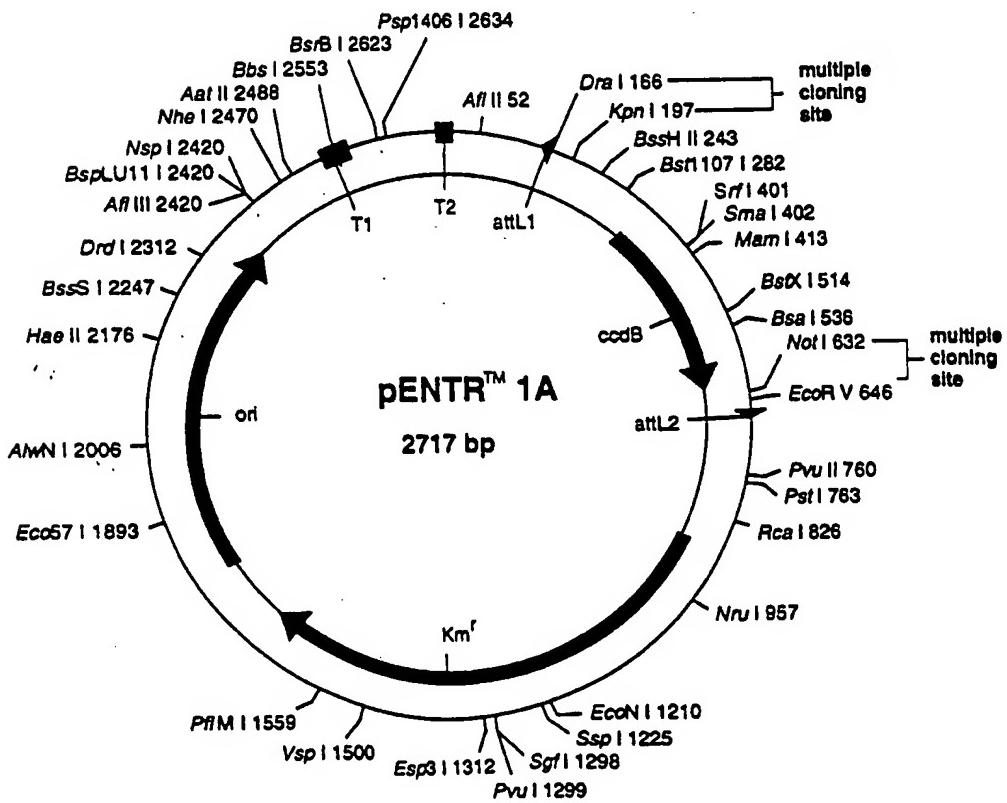
Figure 9

10/260

Figure 10A: Cloning sites of the Entry Vector pENTR1A (reading frame A)

$\underline{Dra\ I}$ $\underline{Xmn\ I}$ $\underline{Sal\ I}$ $\underline{BamH\ I}$ $\underline{Kpn\ I}$ $\underline{EcoR\ I}$
 ACT TTG TAC AAA AAA GCA GGC TTT AAA GGA ACC AAT TCA GTC GAC TGG ATC CGG TAC CGA ATT C
 TGA AAC ATG TTT TTT CGT CCG AAA TTT CCT TGG TTA AGT CAG CTG ACC TAG GCC ATG GCT TAA G
 thr leu tyr lys lys ala gly phe lys gly thr asn ser val asp trp ile arg tyr arg ile

$\underline{EcoR\ I}$ $\underline{Not\ I}$ $\underline{Xba\ I}$ $\underline{EcoR\ V}$
 --- [ccdB gene] --- GIAAT TCG CCG CCG CAC TCG AGA TAT CTA GAC CCA GCT TTC TTG TAC AAA
 C TTA AGC GCC GGC GTG AGC TCT ATA GAT CTG GGT CCA AAG AAC ATG TTT



11/240

pENTR1A 2717 bp

<u>Base Nos.</u>	<u>Gene Encoded</u>
67..166	attL1
321..626	ccdB
655..754	attL2
877..1686	KmR
1791..2364	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
 61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTGCA ACAAAATTGAT
 121 AAGCAATGCT TTTTTATAAT GCCAACTTG TACAAAAAAG CAGGCTTAA AGGAACCAAT
 181 TCAGTCGACT GGATCCGGTA CGAATTTCGC TTACTAAAAG CCAGATAACA GTATCGCTAT
 241 TTGCGCGCTG ATTTTTGCGG TATAAGAATA TATACTGATA TGTATACCG AAGTATGTCA
 301 AAAAGAGGTG TGCTTCTAGA ATGCAGTTA AGGTTTACAC CTATAAAAAGA GAGAGCCGTT
 361 ATCGTCTGTT TGTGGATGTA CAGAGTGATA TTATTGACAC GCCCAGGGCGA CGGATAGTGA
 421 TCCCCCTGGC CAGTCACCGT CTGCTGTCAG ATAAAGTCTC CCGTGAACCT TACCCGGTGG
 481 TGCAATATCGG GGATGAAAGC TGGCGCATGA TGACCACCGA TATGGCCAGT GTGCCGGTCT
 541 CCGTTATCGG GGAAGAAGTG GCTGATCTCA GCACCGCGA AAATGACATC AAAAACGCCA
 601 TTAACCTGAT GTTCTGGGGA ATATAGAATT CGCGGCCGCA CTCGAGATAT CTAGACCCAG
 661 CTTCTTGTGTA CAAAGTTGGC ATTATAAGAA AGCATTGCTT ATCAATTGTT TGCAACGAAC
 721 AGGTCACTAT CAGTCAAAAT AAAATCATTG TTTGCCATCC AGCTGCAGT CTGGCCGGTGT
 781 TCTCAAAATC TCTGATGTTA CATTGACAA GATAAAAATA TATCATCATG AACAAATAAAA
 841 CTGTCTGCTT ACATAAACAG TAATACAGG GGTGTTATGA GCCATATTCA ACGGGAAACG
 901 TCGAGGCCGC GATTAAATC CAACATGGAT GCTGATTATAT ATGGGTATAA ATGGGCTCGC
 961 GATAATGTCG GGCAATCAGG TGCACAAATC TATCGCTTGT ATGGGAAGCC CGATGCGCCA
 1021 GAGTTGTTTC TGAAACATGG CAAAGGTAGC GTGCCAATG ATGTTACAGA TGAGATGGTC
 1081 AGACTAAACT GGCTGACGGA ATTTATGCTT CTCCGACCA TCAAGCATT TATCCGTACT
 1141 CCTGATGATG CATGGTTACT CACCACTGCG ATCCCCGGAA AACACGATT CCAGGTATTA
 1201 GAAGAATATC CTGATTCAAGG TGAAAATATT GTTGATGCGC TGGCAGTGT CCTGCCCGG
 1261 TTGCAATTGCA TTCTGTTTG TAATTGCTT TTTAACAGCG ATCGCTATT TCGTCTCGCT
 1321 CAGGCCGAAT CACGAATGAA TAACGGTTTG GTTGATGCGA GTGATTTGA TGACGAGCGT
 1381 AATGGCTGGC CTGTTGAACA AGTCTGGAAA GAAATGCATA AACCTTTGCC ATTCTCACCG
 1441 GATTCACTCG TCACATCAGG TGATTCTCA CTTGATAACC TTATTTTGA CGAGGGAAA
 1501 TTAATAGGTT GTATTGATGT TGGACGAGTC GGAATCGCAG ACCGATACCA GGATCTGCC
 1561 ATCCTATGGA ACTGCCCTGG TGAGTTTCTC CCTTCATTAC AGAAACGGCT TTTCAAAAAA
 1621 TATGGTATTG ATAATCCTGA TATGAATAA TTGCACTTTC ATTTGATGCT CGATGAGTTT
 1681 TTCTAATCAG AATTGGTTAA TTGGTTGTA CATTATTCAAGG ATTTGGGCCCGG GTTCCACTGA
 1741 GCGTCAGACC CCGTAGAAAA GATCAAAGGA TCTTCCTGAG ATCCTTTTT TCTGCCCGTA
 1801 ATCTGCTGCT TGCAAAACAAA AAAACACCG CTACCAAGCGG TGGTTTGTGTT GCCGGATCAA
 1861 GAGCTACCAA CTCTTTTCC GAAGGTAACG GGCTTCAGCA GAGCGCAGAT ACCAAATACT
 1921 GTTCTCTAG TGTAGCCGTA GTTACGCCAC CACTTCAGA ACTCTGTAGC ACCGCCCTACA
 1981 TACCTCGCTC TGCTAATCCT GTTACCAAGTG GCTGCTGCCA GTGGCGATAA GTCGTGTCTT
 2041 ACCGGGTTGG ACTCAAGACG ATAGTTACCG GATAAGGCAG AGCGGTCGGG CTGAACGGGG
 2101 GTTCTCGCA CACAGCCAG CTTGGAGCGA AGCACCTACA CCGAAGCTGAG ATACCTACAG
 2161 CGTGAGCTAT GAGAAAGCGC CACGCTTCCC GAAGGGAGAA AGGCGGACAG GTATCCGGTA
 2221 AGCGGCAGGG TCGGAACAGG AGAGCGCAGG AGGGAGCTTC CAGGGGGAAA CGCCTGGTAT
 2281 CTTTATAGTC CTGTCGGGTT TCGCACCTC TGACTTGAGC GTGATTTTT GTGATGCTCG
 2341 TCAGGGGGGC GGAGCCTATG GAAAAACGCC AGCAACCGGG CCTTTTTACG GTTCCGGCC
 2401 TTTGCTGGC CTTTCTCA CATGTTCTT CTCGCTTAT CCCCTGATTC TGTGGATAAC
 2461 CGTATTACCG CTAGCATGGA TCTCGGGAC GTCTAACTAC TAAGCGAGAG TAGGAAACTG
 2521 CCAGGCATCA AATAAAACGA AAGGCTCAGT CGGAAGACTG GGCCTTTCGT TTTATCTGTT
 2581 GTTGTGCGGT GAACGCTCTC CTGAGTAGGA CAAATCCGCC GGGAGCGGAT TTGAACGTTG
 2641 TGAAGCAACG GCCCCGGAGGG TGGCGGGCAG GACGCCCGCC ATAAACTGCC AGGCATCAA
 2701 CTAAGCAGAA GGCCATC

FIGURE 10B

12/24/00

Figure 11A: Cloning Sites of the Entry Vector pENTR2B (reading frame B)

Int attL1 EheI XmnI SalI BamHI

TIG TAC AAA AAA GCA GGC TGG CGC CGG AAC CAA TTC AGT CGA CTG |GAT CCG
 AAC ATG TTT| TTT CGT CCG ACC GCG GCC TTG GTT AAG TCA GCT|GAC CTA GSC

↓ ↓ ↓ ↓

Leu Tyr Lys Lys Ala Gly Trp Arg Arg Asn Gln Phe Ser Arg Leu Asp Pro

KpnI EcoRI EcoRI NotI XbaI EcoRV XbaI

GTA CCG |AAT TC- ccdB --G|AAT TCG CGG CCG CAC |TCG AGA TAT|CTA GAC CCA
 CAT GGC TTA AG C TTA AGC GCC GGC GTG AGC TCT ATA GAT CTG GGT

↓ ↓ ↓ ↓ ↓

Val Pro Asn Asn Ser Arg Pro His Ser Arg Tyr Leu Asp Pro

Int attL2

GCT TTC TTG TAC AAA G
CGA AAG AAC ATG TTT C

Ala Phe Leu Tyr Lys

pENTR2B 2718 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
322..627	ccdB
656..755	attL2
878..1687	KmR
1792..2365	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
 61 GGGCCCCAAA TAATGATTG ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAAATTGAT
 121 AAGCAATGCT TTTTATAAT GCCAACCTTG TACAAAAAAG CAGGCTGGCG CGGAAACCAA
 181 TTCAGTCGAC TGGATCCGGT ACCGAATTCTG CTTACTAAAAA GCCAGATAAC AGTATGCGTA
 241 TTTGCGCGCT GATTTTTGCG GTATAAGAAT ATATACGTGAT ATGTATAACCC GAAGTATGTC
 301 AAAAAGAGGT GTGCTTCTAG AATGCAAGTTT AAGGTTTACA CCTATAAAAG AGAGAGCCGT
 361 TATCGTCTGT TTGTGGATGT ACAGAGTGT ATTATTGACA CGCCCGGGCG ACGGATGGTG
 421 ATCCCCCTGG CCAGTGCACG TCTGCTGTCA GATAAAGTCT CCCGTGAACCT TTACCCGGTG
 481 GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCACCG ATATGGCCAG TGTGCCGGTC
 541 TCCGTTATCG GGGAAAGAAGT GGCTGATCTC AGCCCACCGCG AAAATGACAT CAAAACGCC
 601 ATTAACCTGA TGTTCTGGG AATATAGAAT TCGCGGCGC ACTCGAGATA TCTAGACCCA
 661 GCTTTCTTGT ACAAAGTTGG CATTATAAGA AAGCATTGCT TATCAATTG TTGCAACGAA
 721 CAGGTCACTA TCAGTCAAA TAAAATCATT ATTGCCATC CAGCTGCAGC TCTGGCCCGT
 781 GTCTCAAAAT CTCTGATGTT ACATTGCAAGA AGATAAAAAT ATATCATCAT GAACAATAAA
 841 ACTGTCTGCT TACATAAACAGA GTAATAGAAG GGGTGTATG AGCCATATTC AACGGGAAAC
 901 GTCGAGGGCCG CGATTAATT CCAACATGGA TGCGTATTAA TATGGGTATA AATGGGCTCG
 961 CGATAATGTC GGGCAATCAG GTGCGACAAT CTATCGCTTG TATGGGAAGC CCGATGCGCC
 1021 AGAGTGTGTT CTGAAACATG GCAAAGGTAG CGTTGCCAAT GATGTTACAG ATGAGATGGT
 1081 CAGACTAACAC TGGCTGACGG AATTATGCC TCTTCCGACC ATCAAGCATT TTATCCGTAC
 1141 TCCTGATGAT GCATGGTTAC TCACCACTGC GATCCCCGGA AAAACAGCAT TCCAGGTATT
 1201 AGAGAAATAT CCTGATTCAAG GTGAAAATAT TGTTGATGCG CTGGCAGTGT TCCCGCCCG
 1261 GTTGCATTG ATTCCTGTT GTATTGTC TTTAACAGC GATCGGTAT TTCTGCTCGC
 1321 TCAGGCGCAA TCACGAATGA ATAACGGTTT GGTTGATGCG AGTGATTTTG ATGACGAGCG
 1381 TAATGGCTGG CCTGTTGAAC AAGTCTGGAA AGAAAATGCAT AAACTTTGC CATTCTCAC
 1441 GGATTCACTC GTCACTCATG GTGATTCTC ACTTGATAAC CTTATTTTG ACGAGGGAA
 1501 ATTAATAGGT TGTATTGATG TTGGACGAGT CGGAATCGCA GACCAGTAC AGGATCTTGC
 1561 CATCCTATGG AACTGCCTCG GTGAGTTTC TCCTTCATTA CAGAAACGGC TTTTCAAAA
 1621 ATATGGTATT GATAATCTG ATATGAATAA ATTGCAGTTT CATTGATGC TCGATGAGTT
 1681 TTTCTAATCA GAATTGGTTA ATGGGGTGA ACATTATTCA GATTGGGCC CGTTCCACTG
 1741 AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT
 1801 AATCTGCTGC TTGCAAACAA AAAACCACCG GCTACCAGCG GTGGTTGTT TGCGGGATCA
 1861 AGAGCTACCA ACTCTTTTC CGAAGGTAAC TGCGTTCAGC AGAGCGAGA TACCAAATAC
 1921 TGTTCTTCTA GTGTAGCCG AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC
 1981 ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT
 2041 TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG
 2101 GGGTTCTGTC ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACCTGA GATAACCTACA
 2161 GCCTGAGCTA TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT
 2221 AAGCGGCAGG GTCCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA ACCGCTGGTA
 2281 TCTTTATAGT CCTGTCGGGT TTCGCCCCCT CTGACTTTGAG CGTCGATTTT TGTGATGCTC
 2341 GTCAGGGGGG CGGAGCCTAT GGAAAAAACGC CAGCAACCGCG GCCTTTTAC GGTTCTGGC
 2401 CTTTGCTGG CCTTTTGCTC ACATGTTCTT CCTGCGTTA TCCCCCTGATT CTGTTGGATAA
 2461 CCGTATTACG GCTAGCATGG ATCTCGGGGA CGTCTAACTA CTAAGCGAGA GTAGGAACT
 2521 GCCAGGCATC AAATAAAAAG AAAGGCTAG TCGGAAGACT GGGCCTTTCG TTTTATCTGT
 2581 TGTTGTCGG TGAACGCTCT CCTGAGTAGG ACAAAATCCGC CGGGAGCGGA TTTGAACGTT
 2641 GTGAAGCAAC GGCCCCGGAGG GTGGCGGGCA GGACGCCGC CATAAACTGC CAGGCATCAA
 2701 ACTAAGCAGA AGGCCATC

FIGURE 1B

14/260

Figure 12A: Cloning Sites of the Entry Vector pENTR3C (reading frame C) ...

Int	attL1	DraI	XbaI	SalI	BamHI
-----	-------	------	------	------	-------

TTG TAC AAA AAA GCA GGC TCT TAA AAG GAA CCA ATT CAG TCG ACT GGA TCC GGT
 AAC ATG TTT TTT CGT CCG AGA AAT TTC CTT GGT TAA GTC AGC TGA CCT AGG CCA
 ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
 Leu Tyr Lys Lys Ala Gly Ser Leu Lys Glu Pro Ile Gln Ser Thr Gly Ser Gly

KpnI	EcoRI	PvuI	EcoRI	NotI	XbaI	EcoRV	XbaI
------	-------	------	-------	------	------	-------	------

ACC GAA TTC GAT GGC -- ccdB -- G AAT TCG CGG CCG CAC TCG AGA TAT CTA
 TGG CTT AAG CTA GCG ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
 Thr Glu Phe Asn Ser Arg Pro His Ser Arg Tyr Leu

attL2	Int
-------	-----

GAC CCA GCT TTC TTG TAC AAA G
 CTG GGT CGA AAG AAC ATG TTT C
 ↓
 Asp Pro Ala Phe Leu Tyr Lys

15/240

pENTR3C 2723 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
327..632	ccdB
661..760	attL2
883..1692	KmR
1797..2370	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
 61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT
 121 AAGCAATGCT TTTTTATAAT GCCAACCTTG TACAAAAAAG CAGGCTCTT AAAGGAACCA
 181 ATTCACTCGA CTGGATCCGG TACCGAATTG GATCGCTTAC TAAAAGCCAG ATAACAGTAT
 241 GCGTATTGTC GCGCTGATTT TTGCGGTATA AGAATATATA CTGATATGTA TACCCGAAGT
 301 ATGTCAAAAAA GAGGTGTGCT TCTAGAATGCA AGTTTAAGGT TTACACCTAT AAAAGAGAGA
 361 GCCGTTATCG TCTGTTTG TGATGTACAGA GTGATATTAT TGACACGGCC GGGCGACGGA
 421 TGGTGTATCCC CCTGGCCAGT GCACCTCTGC TGTCAGATAA AGTCTCCGT GAACCTTACCC
 481 CGGTGGTGC TATCGGGGAT GAAAGCTGGC GCATGATGAC CACCGATATG GCCAGTGTGC
 541 CGGTCTCCGT TATCGGGGAA GAAGTGGCTG ATCTCAGCCA CCGCGAAAAT GACATCAAAA
 601 ACGCCATTAA CCTGATGTTG TGCGGAAATAT AGAATTCGCG GCGCAGCTG AGATATCTAG
 661 ACCCAGCTT CTTGTACAAA GTTGGCATTAA TAAGAAAGCA TTGCTTATCA ATTTGTTGCA
 721 ACGAACAGGT CACTATCAGT CAAAATAAAA TCATTATTTG CCATCCAGCT GCAGCTCTGG
 781 CCCGTGTCTC AAAATCTCTG ATGTTACATT GCACAAGATA AAAATATATC ATCATGAACA
 841 ATAAAACGTG CTGCTTACAT AAACAGTAAT ACAAGGGGTG TTATGAGGCC TATTCAACGG
 901 GAAACGTGCA GGCGCGATT AAATTCCAAC ATGGATGCTG ATTTATATGG GTATAAAATGG
 961 GCTCGCGATA ATGTCGGCA ATCAGGTGCG ACAATCTATC GCTTGTATGG GAAGCCCGAT
 1021 GCGCCAGAGT TGTTCTGAA ACATGGCAAAG GTAGCGTTG CCAATGATGT TACAGATGAG
 1081 ATGGTCAGAC TAAACTGGCT GACGGAATTG ATGCTCTTC CGACCATCAA GCATTTTATC
 1141 CGTACTCCTG ATGATGCATG GTTACTCACC ACTGCGATCC CCGGAAAAAC AGCATTCCAG
 1201 GTATTAGAAG AATATCCTGA TTCAGGTGAA AATATTGTG ATGCGCTGGC AGTGTTCCTG
 1261 CGCCGGTTGC ATTGATTCC TGTTTGTAAAT TGTCCTTTA ACAGCGATCG CGTATTTCGT
 1321 CTCGCTCAGG CGCAATCACG AATGAATAAC GGTTTGGTTG ATGCGAGTGA TTTGATGAC
 1381 GAGCGTAATG GCTGGCCTGT TGAACAAGTC TGAAAGAAA TGCATAAACT TTTGCCATTC
 1441 TCACCGGATT CAGTCGTAC TCATGGTGTAT TTCTCACTTG ATAACCTTAT TTTGACGAG
 1501 GGGAAATTAA TAGGTTGTAT TGATGTTGGA CGAGTCGGAA TCGCAGACCG ATACCAGGAT
 1561 CTTGCCATCC TATGGAACCTG CCTCGGTGAG TTTTCTCCTT CATTACAGAA ACGGCTTTT
 1621 CAAAAATATG GTATTGATAA TCCTGATATG AATAAAATTG AGTTTCATTT GATGCTCGAT
 1681 GAGTTTTCT AATCAGAATT GGTTAATTGG TTGTAACATT ATTCAAGATTG GGCCCCGTTC
 1741 CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTCTG
 1801 CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCGCTAC CAGCGGGTGT TTGTTGCG
 1861 GATCAAGAGC TACCAACTCT TTTCCGAAG GTAACTGGCT TCAGCAGAGC GCAGATAACCA
 1921 AATACTGTTC TCTCTAGTGA GCCGTAGTTA GGCCACACT TCAAGAACTC TGTAGCACCG
 1981 CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG
 2041 TGTCTTACCG GGTTGGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA
 2101 ACGGGGGGTT CGTGACACCA GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC
 2161 CTACAGCGT AGCTATGAGA AAGGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT
 2221 CCGGTAAGCG GCAGGGTCGG AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC
 2281 TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTGTGA
 2341 TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA ACGGGCCCTT TTTACGGTTC
 2401 CTGGCTTTT GCTGGCTTT TGCTCACATG TTCTTCTG CGTTATCCCC TGATTCTGTG
 2461 GATAACCGTA TTACCGCTAG CATGGATCTC GGGGACGTCT AACTACTAAG CGAGAGTAGG
 2521 GAACTGCCAG GCATCAAATA AAACGAAAGG CTCAGTCGGAA AGACTGGGCC TTTCGTTTA
 2581 TCTGTTGTTT GTCGGTGAAC GCTCTCCTGA GTAGGACAAA TCCGCGGGGA GCGGATTGAA
 2641 ACGTTGTGAA GCAACGGCCC GGAGGGTGGC GGGCAGGACG CCCGCCATAA ACTGCCAGGC
 2701 ATCAAACCTAA GCAGAAGGCC ATC

FIGURE 12B

16/240

Figure 13A: Cloning Sites of the Entry Vector pENTR4

Int attL1 NcoI Kozak XmnI SalI BamHI
TTG TAC AAA AAA GCA GGC TCC ACC ATG GGA ACC AAT TCA GTC GAC TGG ATC CCG
AAC ATG TTT TTT CGT CCG AGG TGG TAC CCT TGG TTA AGT CAG CTG ACC TAG GCC
Leu Tyr Lys Lys Ala Gly Ser Thr Met Gly Thr Asn Ser Val Asp Trp Ile Arg

Int attL2

TTC TTG TAC AAA G
AAG AAC ATG TTT C

17/240

pENTR4 2720 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
324..629	ccdB
658..757	attL2
880..1689	KmR
1794..2367	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
 61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGAACCT GTTCGTGCA ACAATTGAT
 121 AAGCAATGCT TTTTTATAAT GCCAACTTTG TACAAAAAAG CAGGCTCCAC CATGGGAACC
 181 AATTCACTCG ACTGGATCCG GTACCGAATT CGCTTACTAA AAGCCAGATA ACAGTATGCG
 241 TATTTCGCGC CTGATTTTTG CCGTATAAGA ATATATACTG ATATGTATAC CCGAAGTATG
 301 TCACAAAGAG GTGTGCTTCT AGAACGAGT TTAAGGTTA CACCTATAAA AGAGAGAGCC
 361 GTTATCGTCT GTTTGTGGAT GTACAGAGTG ATATTATTGA CACGCCCGGG CGACGGATGG
 421 TGATCCCCCT GGCCAGTGC CGTCTGCTGT CAGATAAAAGT CTCCCGTCAA CTTTACCCGG
 481 TGGTGCATAT CGGGGATGAA AGCTGGCGCA TGATGACCAC CGATATGGCC AGTGTGCCGG
 541 TCTCCGTTAT CGGGGAAGAA GTGGCTGATC TCAGGCCACCG CGAAAATGAC ATCAAAAACG
 601 CCATTAACCT GATGTTCTGG GGAATATAGA ATTGCGGGCC GCACTCGAGA TATCTAGACC
 661 CAGCTTCTT GTACAAAGTT GGCATTATAA GAAAGCATTG CTTATCAATT TGTTGCAACG
 721 AACAGGTCAC TATCAGTCAA AATAAAATCA TTATTTGCCA TCCAGCTGCA GCTCTGGCC
 781 GTGTCTCAA ATCTCTGATG TTACATTGCA CAAGATAAAA ATATATCATC ATGAACAATA
 841 AAACGTCTG CTTACATAAA CAGTAATACA AGGGGTGTTA TGAGCCATAT TCAACGGAA
 901 ACGTGAGGC CGCGATTAAA TTCCAACATG GATGCTGATT TATATGGGTAA TAAATGGGCT
 961 CGCGATAATG TCGGGCAATC AGGTGCGACA ATCTATCGCT TGTATGGGAA GCCCGATGCG
 1021 CCAGAGTTGT TTCTGAAACA TGGCAAAGGT AGCGTTGCCA ATGATGTTAC AGATGAGATG
 1081 GTCAGACTAA ACTGGCTGAC GGAATTATG CCTCTTCCGA CCATCAAGCA TTTTATCCGT
 1141 ACTCCTGGTG ATGCATGGTT ACTCACCACT GCGATCCCCG GAAAACAGC ATTCCAGGTA
 1201 TTAGAAGAAT ATCCTGATTC AGGTGAAAAT ATTGTTGATG CGCTGGCAGT GTTCTGGCC
 1261 CGGTTGCATT CGATTCCCTGT TTGTAATTGT CCTTTAACA GCGATCGCGT ATTTCTGCTC
 1321 GCTCAGGCGC AATCACGAAT GAATAACGGT TTGGTTGATG CGAGTGATTT TGATGACGAG
 1381 CGTAATGGCT GGCCTGTTGA ACAAGTCTGG AAAGAAATGC ATAAACCTTT GCCATTCTCA
 1441 CCGGATTTCAG TCGTCACTCA TGGTGAATTTC TCACTTGATA ACCTTATTGTT TGACGAGGGG
 1501 AAAAAATAG GTTGATTTGA TGTTGGACGA GTCGGAATCG CAGACCGATA CCAGGATCTT
 1561 GCCATCCTAT GGAACTGCCT CGGTGAGTTT TCTCCTTCAT TACAGAAAAG GCTTTTTCAA
 1621 AAATATGGTA TTGATAATTC TGATATGAAT AAATTGCACT TTCAATTGAT GCTCGATGAG
 1681 TTTTCTAAT CAGAATTGTT TAATTGGTTG TAAACATTATT CAGATTGGC CCCGTTCCAC
 1741 TGAGCGTCAG ACCCGTAGA AAAGATCAAA GGATCTTCTT GAGATCCTTT TTTCTGGCC
 1801 GTAATCTGCT GCTTGAAAC AAAAAGACCA CGGCTACCAAG CGGTGGTTTG TTTGCCGGAT
 1861 CAAGAGCTAC CAAACTTTT TCCGAAGGTA ACTGGCTTCA GCAGAGCGCA GATACCAAAT
 1921 ACTGTTCTTC TAGTGTAGCC GTAGTTAGGC CACCACTTCA AGAACTCTGT AGCACCGCCT
 1981 ACATACCTCG CTCCTGCTAAT CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT
 2041 CTTACCGGGT TGGACTCAAG ACGATAGTTA CCGGATAAGG CGCAGCGTC GGGCTGAACG
 2101 GGGGGTTCGT GCACACAGCC CAGCTGGAG CGAACGACCT ACACCGAAT GAGATAACCTA
 2161 CAGCGTGAGC TATGAGAAAG CGCCACGCTT CCCGAAGGGGA GAAAGGCGGA CAGGTATCCG
 2221 GTAAGCGGCA GGGTCGGAAC AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCCCTGG
 2281 TATCTTTATA GTCTCTGCGG GTTTGCCAC CTCTGACTTG AGCGTCGATT TTTGTGATGC
 2341 TCGTCAGGGG GGGGGAGCCT ATGGAAAAAC GCCAGCAACG CGGCCTTTTT ACGGTTCTG
 2401 GCCTTTGCT GGCCTTTGC TCACATGTTT TTCTCTGCGT TATCCCCTGA TTCTGTGGAT
 2461 AACCGTATTA CCGCTAGCAT GGATCTCGGG GACGTCTAAC TACTAAGCGA GAGTAGGGAA
 2521 CTGCCAGGCA TCAAATAAAA CGAAAGGCTC AGTCGGAAGA CTGGGGCTTT CGTTTTATCT
 2581 GTTGTGTTGTC GGTGAACGCT CTCCCTGAGTA GGACAAATCC GCGGGGAGCG GATTTGAACG
 2641 TTGTGAAGCA ACGGCCCCGA GGGTGGCGGG CAGGACGCCC GCCATAAACT GCCAGGCATC
 2701 AACTAAGCA GAAGGCCATC

FIGURE 13B

18/240

Figure 14A: Cloning sites of the Entry Vector pEMRS

Int att-L1 Nde I Kpn I Sal I
 Tgg tac aaa aaa gca ggc ttat ctt atg gga tcc aat tca atc
tac atg ttt ttt cgt ccg ana gta tac eet egg tta agt cag
 Leu Tyr Lys Lys Ala Gly Phe His Met Gly Thr Asn Ser Val

Bam H I Kpn I Eco RI Eco RI
 gac tgg atc cgg tac cgt att cgc --- Death --- aat att cgc
cgg acc tag ggc atg get taa gcg --- (ccdB) --- tct taa geg.
 Asp Trp Ile Arg Tyr Arg Ile

Nhe I Xba I Eco RI Xba Int att-L2
 bgo cgc att cga gat atc tag acc cag ctt tcc xgg aca adg
ccg cgg tga get cta tag atc tgg gtc gaa aca aca tct tcc ---

19/260

pENTR5 2720 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
324..629	ccdB
658..757	attL2
880..1689	KmR
1794..2367	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
 61 GGGCCCCAAA TAATGATTTC ATTTTGACTG ATAGTGCCT GTTCGTTGCA ACAAAATTGAT
 121 AAGCAATGCT TTTTTATAAT GCCAACCTTG TACAAAAAAG CAGGCTTTCA TATGGGAACC
 181 AATTCACTCG ACTGGATCCG GTACCGGATT CGCTTACTAA AAGCCAGATA ACAGTATGCG
 241 TATTGCGCG CTGATTTTG CCGGTATAAGA ATATATACTG ATATGTATAC CCGAAGTATG
 301 TCAAAAAGAG GTGTGCTTCT AGAATGCACT TTAAGGTTA CACCTATAAA AGAGAGAGCC
 361 GTTATCGTCT GTTTGTGGAT GTACAGAGTG ATATATTGTA CACGCCCGGG CGACGGATGG
 421 TGATCCCCCT GGCCAGTGC A CGTCTGCTGT CAGATAAAAGT CTCCCGTCAA CTTTACCCGG
 481 TGGTGCATAT CGGGGATGAA AGCTGGCGCA TGATGACCAC CGATATGGCC AGTGTGCCGG
 541 TCTCCGTTAT CGGGGAAGAA GTGGCTGATC TCAGCCACCG CGAAAATGAC ATCAAAAACG
 601 CCATTAACCT GATGTTCTGG GGAATATAGA ATTCCGGGCC GCACTCGAGA TATCTAGACC
 661 CAGCTTTCTT GTACAAAGTT GGCATTATAA GAAAGCATTG CTTATCAATT TGTTGCAACG
 721 AACAGGTCAC TATCAGTCAA AATAAAATCA TTATTTGCCA TCCAGCTGCA GCTCTGGCCC
 781 GTGTCTAAA ATCTCTGATG TTACATTGCA CAAGATAAAA ATATATCATC ATGAACAATA
 841 AAACGTCTG CTTACATAAA CAGTAATACA AGGGGTGTTA TGAGCCATAT TCAACGGGAA
 901 ACGTGAGGC CGCGATTAAA TTCCAACATG GATGCTGATT TATATGGGTAA TAAATGGGCT
 961 CGCGATAATG TCGGGCAATC AGGTGCGACA ATCTATCGCT TGATGGGAA GCCCGATGCG
 1021 CCAGAGTTGT TTCTGAAACA TGGCAAAGGT AGCGTTGCCA ATGATGTTAC AGATGAGATG
 1081 GTCAGACTAA ACTGGCTGAC GGAATTATG CCTCTTCGG CCATCAAGCA TTTTATCCGT
 1141 ACTCCTGATG ATGCATGGTT ACTCACCACT GCGATCCCCG GAAAACAGC ATTCCAGGTA
 1201 TTAGAAGAAT ATCCTGATTC AGGTGAAAT ATTGTTGATG CGCTGGCAGT GTTCTGCGC
 1261 CGGTTGCATT CGATTCTGT TTGTAATTGT CCTTTTAACA CGCAGTCGCGT ATTTCTGCTC
 1321 GCTCAGGCGC AATCACGAAT GAATAACGGT TTGGGTTGATG CGAGTGATTT TGATGACGAG
 1381 CGTAATGGCT GGCCTGTTGA ACAAGTCTGG AAGAAATGCA ATAAACTTT GCCATTCTCA
 1441 CCGGATTTCAG TCGTCACTCA TGGTGAATTTC TCACTTGATA ACCTTATTG TGACGAGGGG
 1501 AAATTAATAG GTTGTATTGA TGGTGGACGA GTCGGAATCG CAGACCGATA CCAGGATCTT
 1561 GCCATCCTAT GGAACTGCT CGGTGAGTT TCTCCTTCAT TACAGAAACG GCTTTTCAA
 1621 AAATATGGTA TTGATAATCC TGATATGAAT AAATTGCACT TTCATTGAT GCTCGATGAG
 1681 TTTTCTAAT CAGAATTGTTA TAAATTGGTTG TAACTATT CAGATTGGGC CCCGTCCAC
 1741 TGAGCGTCAG ACCCCGTAGA AAAGATCAAA GGATCTTCTT GAGATCCTT TTTCTGCGC
 1801 GTAATCTGCT GCTTGCAAAAC AAAAAGACCA CGGCTTACAG CGGTGGTTG TTTGCCGGAT
 1861 CAAGAGCTAC CAACTCTTT TCCGAAGGTA ACTGGCTTCA GCAGAGCGCA GATACCAAT
 1921 ACTGTTCTTC TAGTGTAGCC GTAGTTAGGC CACCACTTCAGA AGAACTCTGT AGCACCGCCT
 1981 ACATACCTCG CTCTGCTAAT CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT
 2041 CTTACGGGT TGGACTCAAG ACGATAGTTA CCGGATAAGG CGCAGCGTC GGGCTGAACG
 2101 GGGGGTTCGT GCACACAGCC CAGCTTGGAG CGAACGACCT ACACCGAACT GAGATAACCA
 2161 CAGCGTGAGC TATGAGAAAG CGCCACGCTT CCCGAAGGGAA GAAAGGGCGA CAGGTATCCG
 2221 GAGCGGGCA GGGTCGGAAC AGGAGAGCGC AGGAGGGAGC TTCCAGGGGG AAACGCCCTGG
 2281 TATCTTTATA GTCTCTGTCGG GTTTGCCAC CTCTGACTTG AGCGTCGATT TTTGTGATGC
 2341 TCGTCAGGGG GGGGGAGCCT ATGGAAAAC GCCAGCAACG CGGCCTTTT ACGGTTCCGT
 2401 GCCTTTGCT GGCCTTTGC TCACATGTC TTCTCTGCGT TATCCCCCTGA TTCTGTGGAT
 2461 AACCGTATTA CCGCTAGCAT GGATCTCGGG GACGTCTAAC TACTAAGCGA GAGTAGGGAA
 2521 CTGCCAGGCA TCGAATAAAA CGAAAGGCTC AGTCGGAAGA CTGGGCTTT CGTTTATCT
 2581 GTTGTGTC GGTGAACGCT CTCCTGAGTA GGACAAATCC GCGGGGAGCG GATTGAAACG
 2641 TTGTGAAGCA ACGGCCCGGA GGGTGGCGGG CAGGACGCC GCGATAAACT GCCAGGCATC
 2701 AACTAAGCA GAAGGCCATC

Figure 14B

20/240

Figure 1/5A: Cloning sites of the Entry Vector pENTR 6

Int att L1 Sph I Kpn I Sal I
 --- tac aaa aaa gca ggc tgg atg cga acc aat tca gtc
 --- aac aag ttt cgt ccg agg tao get tgg tta agt cag
 Leu Tyr Lys Lys Ala Gly Cys Met Arg Thr Asn Ser Val

BamH I Kpn I EcoRI EcoRI
 gac tgg atc ogg tac cgt att cgo --- Death --- aga att cgc
 cgg acc tag gct atg get taa gag --- (codB) --- tct taa gcg
 Asp Trp Ile Arg Tyr Arg Ile

Not Xba I EcoRI Xba I Int att L2
 ggc cgc act cga gat atc tag acc cag ctt dgr tgg aga dag ---
 cgg ggg tga gct cta tag atc tgg gtc gaa aga aca tgt rcc ---

21/240

pENTR6 2717 bp

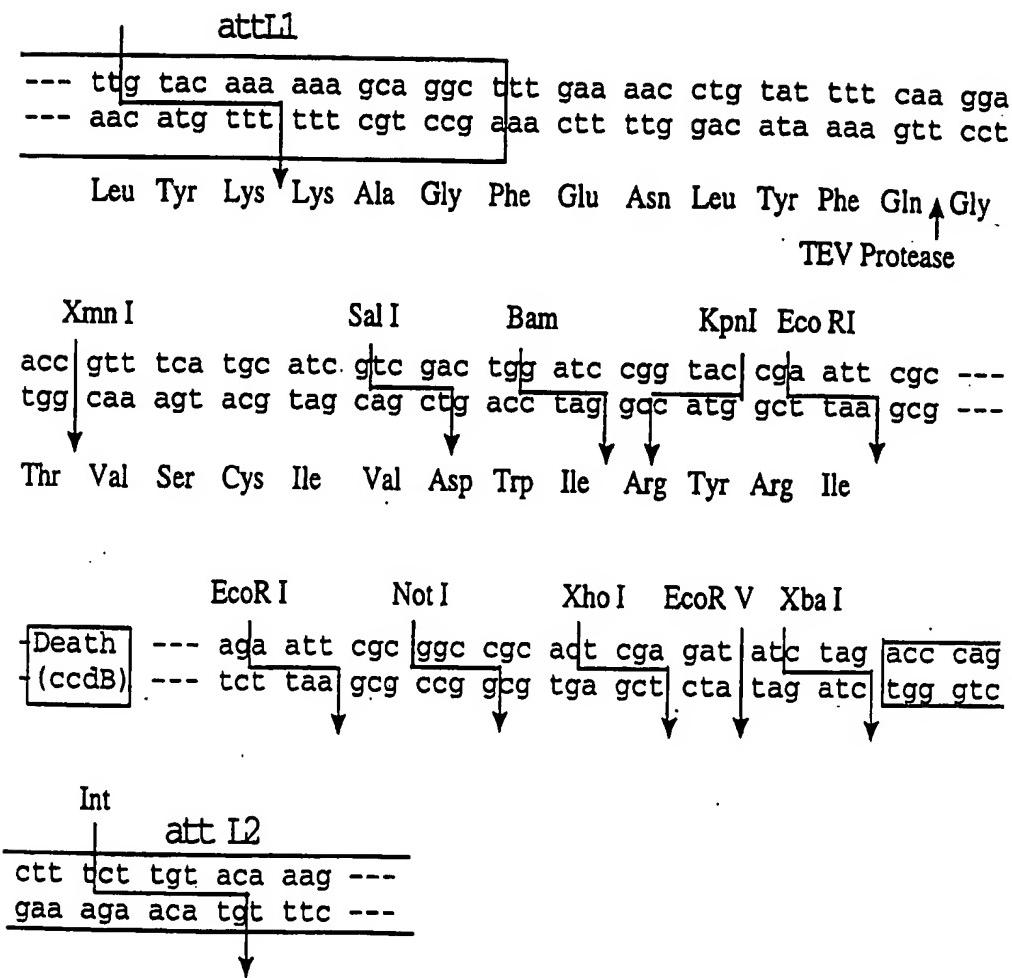
<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
321..626	ccdB
655..754	attL2
877..1686	KmR
1791..2364	ori

1 CTGACGGATG GCCTTTTFTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
 61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAAATTGAT
 121 AAGCAATGCT TTTTTATAAT GCCAACTTG TACAAAAAAG CAGGCTGCAT GCGAACCAAT
 181 TCAGTCGACT GGATCCGGTA CGAATTTCGC TTACTAAAAG CCAGATAACA GTATGCGTAT
 241 TTGCGCGCTG ATTTTTGCGG TATAAGAATA TATACTGATA TGTATACCCG AAGTATGTCA
 301 AAAAGAGGTG TGCTTCTAGA ATGCAGTTA AGGTTTACAC CTATAAAAGA GAGAGCCGTT
 361 ATCGTCTGTT TGTGGATGTA CAGAGTATA TTATTGACAC GCCCCGGCGA CGGATGGTGA
 421 TCCCCCTGGC CAGTGCACGT CTGCTGTCAG ATAAAGTCTC CCGTGAACCT TACCCGGTGG
 481 TGCATATCGG GGATGAAAGC TGGCGCATGA TGACCAACCGA TATGCCAGT GTGCCGGTCT
 541 CCGTTATCGG GGAAGAAGTG GCTGATCTCA GCACCCGCGA AAATGACATC AAAAACGCCA
 601 TTAACCTGAT GTTCTGGGA ATATAGAATT CGCGGCCGCA CTCGAGATAT CTAGACCCAG
 661 CTTCTTGTA CAAAGTTGGC ATTATAAGAA AGCATTGCTT ATCAATTGTT TGCAACGAAAC
 721 AGGTCACTAT CAGTCAAAAT AAAATCATT TTTGCCATCC AGCTGCAGCT CTGGCCCGTG
 781 TCTCAAAATC TCTGATGTTA CATTGCACAA GATAAAAATA TATCATCATG AACATAAAA
 841 CTGTCGCTT ACATAAACAG TAATACAAGG GGTGTTATGA GCCATATTCA ACGGGAAACG
 901 TCGAGGCCGC GATTAAATTG CAACATGGAT GCTGATTAT ATGGGTATAA ATGGGCTCGC
 961 GATAATGTCG GGCAATCAGG TGCGACAATC TATGCTTGT ATGGGAAGCC CGATGCCGCC
 1021 GAGTTGTTTC TGAAACATGG CAAAGGTAGC GTTCCAATG ATGTTACAGA TGAGATGGTC
 1081 AGACTAAACT GGCTGACCGA ATTTATGCC CTTCCGACCA TCAAGCATT TATCCGTACT
 1141 CCTGATGATG CATGGTTACT CACCACTGCG ATCCCCGGAA AAACAGCATT CCAGGTATTA
 1201 GAAGAATATC CTGATTCAAG TGAAAATATT GTTGATGCGC TGGCAGTGTT CCTGCCCGG
 1261 TTGCAATTGCA TTCTGTTTG TAATTGTCCT TTTAACAGCG ATCGCGTATT TCGTCTCGT
 1321 CAGGCGCAAT CACGAATGAA TAACGGTTTG GTTGAATGCGA GTGATTGTTGA TGACGAGCGT
 1381 AATGGCTGGC CTGTTGAACA AGTCTGGAAA GAAATGCATA AACTTTGCC ATTCTCACCG
 1441 GATTCACTCG TCACCATGAG TGATTCTCA CTTGATAACC TTATTTTGA CGAGGGAAA
 1501 TTAATAGGTT GTATTGATGT TGGACGAGTC GGAATCGCAG ACCGATAACCA GGATCTTGC
 1561 ATCCCTATGGA ACTGCCCTGG TGAGTTTCT CTTTCATTAC AGAAACGGCT TTTCAAAAAA
 1621 TATGGTATTG ATAATCTGA TATGAATAAA TTGCACTTGC ATTGATGCT CGATGAGTTT
 1681 TTCTAATCAG AATGGTTAA TTGGTGTGAA CATTATTGAG ATTGGGCCCG GTCGCACTGA
 1741 GCGTCAGACC CCGTAGAAAA GATCAAAGGA TCTTCTTGAG ATCTTTTTT TCTGCCCGTA
 1801 ATCTGCTGCT TGCAAAACAAA AAAACCCACCG CTACCGCGG TGGTTGTTT GCCGGATCAA
 1861 GAGCTACCAA CTCTTTTCTG GAAGGTAACG GGCTTCAGCA GAGCGCAGAT ACCAAATACT
 1921 GTTCTTCTAG TGTAGCCGA GTTACGCCAC CACTTCAAGA ACTCTGTAGC ACCGCCCTACA
 1981 TACCTCGCTC TGCTAATCCT GTTACCACTG GCTGCTGCCA GTGGCGATAA GTCGTCCTT
 2041 ACCGGGTTGG ACTCAAGAGC ATAGTTACCG GATAAGGCAG AGCGGTGGG CTGAACGGGG
 2101 GGTTCTGCA CACAGCCAG CTTGGAGCGA AGGACCTACA CCGAACTGAG ATACCTACAG
 2161 CGTGAGCTAT GAGAAAGCCG CACGCTTCCC GAAGGGAGAA AGGCGGACAG GTATCCGTA
 2221 AGCGGCAGGG TCGGAACAGG AGAGCGCACG AGGGAGCTTC CAGGGGAAA CGCCCTGGTAT
 2281 CTTTATAGTC CTGTCGGGTT TCGCCACCTC TGACITTGAGC GTCGATTGTT GTGATGCTCG
 2341 TCAGGGGGGC GGAGCCTATG GAAAAACGCC AGCAACCGGG CCTTTTTACG GTTCTGGCC
 2401 TTTTGCTGGC CTTTGCTCA CATGTTCTT CCTGCGTTAT CCCCTGATTG TGTGGATAAC
 2461 CGTATTACCG CTAGCATGGA TCTCGGGGAC GTCTAACTAC TAAGCGAGAG TAGGAAACTG
 2521 CCAGGCATCA AATAAAACGA AAGGCTCAGT CGGAAGACTG GGCCTTTGCG TTTATCTGTT
 2581 GTTGTGCGGT GAACGCTCTC CTGAGTAGGA CAAATCCGCC GGGAGCGGAT TTGAACGTTG
 2641 TGAAGCAACG GCCCCGGAGGG TGGCGGGCAG GACGCCCGCC ATAAACTGCC AGGCATCAAA
 2701 CTAAGCAGAA GGCCATC

Figure 15B

22/240

Figure 16A: Cloning sites of the Entry Vector pENTR7



23/240

pENTR7 2738 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
342..647	ccdB
676..775	attL2
898..1707	KmR
1812..2385	ori

1 CTGACGGATG GCCTTTTGC GTTTCACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
 61 GGGCCCCAAA TAATGATTG ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAAATTGAT
 121 AAGCAATGCT TTTTATAAT GCCAACTTG TACAAAAAAG CAGGCTTGA AAACCTGTAT
 181 TTTCAAGGAA CCGTTTCATG CATCGTCGAC TGGAATCCGGT ACCGAATTG CTTACTAAA
 241 GCCAGATAAAC AGTATGCGTA TTTGCGCGT GATTTTGCG GTATAAGAAT ATATACTGAT
 301 ATGTATAACCC GAAGATGTC AAAAGAGGTT GTGCTTCTAG AATGCAGTT AAGGTTTACA
 361 CCTATAAAAAG AGAGAGCCGT TATCGTCTGT TTGTTGGATGT ACAGAGTGT ATTATTGACA
 421 CGCCCGGGCG ACGGATAGTG ATCCCCCTGG CCAGTGCACG TCTGCTGTC GATAAAGTCT
 481 CCCGTGAACG TTACCCGGTG GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCACCG
 541 ATATGGCCAG TGTGCCGGTC TCCGTATCG GGGAAAGAAGT GGCTGATCTC AGCCACCGCG
 601 AAAATGACAT CAAAAACGCC ATTAACCTGA TTGTTCTGGGG AATATAGAAT TCGCGGCCGC
 661 ACTCGAGATA TCTAGACCCA GCTTTCTTGT ACAAAGTTGG CATTATAAGA AAGCATTGCT
 721 TATCAATTG TTGCAACGAA CAGGTCACTA TCAGTCAAA TAAAATCATT ATTTGCCATC
 781 CAGCTGCAGC TCTGGCCCGT GTCTCAAAAT CTCTGATGTT ACATTGCACA AGATAAAAAT
 841 ATATCATCAT GAACAATAAA ACTGCTCTGT TACATAAAACA GTAATACAAG GGGTGTATG
 901 AGCCATATTG AACGGGAAAC GTCGAGGCCG CGATTTAAATT CCAACATGGG TGCTGATTAA
 961 TATGGGTATA AATGGGCTCG CGATAATGTC GGGCAATCAG GTGCGACAAT CTATCGCTTG
 1021 TATGGGAAGC CCGATGCGCC AGAGTTGTTT CTGAAACATG GCAAAGGTAG CGTTGCCAAT
 1081 GATGTTACAG ATGAGATGGT CAGACTAAAC TGGCTGACGG AATTTATGCC TCTTCCGACC
 1141 ATCAAGCATT TTATCCGTAC TCCTGATGAT GCATGGTTAC TCACCACTGC GATCCCCGGG
 1201 AAAACAGCAT TCCAGGTATT AGAAGAATAT CCTGATTTCAG GTGAAAATAT TGTTGATGCG
 1261 CTGGCAGTGT TCCCTGCCCG GTTGCATTG ATTCCCTGTTT GTAATTGTC TTTTAACAGC
 1321 GATCGCGTAT TTCGTCTGC TCAGGCCAA TCACGAATGA ATAACGGTTT GGTTGATGCC
 1381 AGTGAATTG ATGACGAGCG TAATGGCTGG CCTGTTGAAC AAGTCTGGAA AGAAATGCAT
 1441 AAACTTTGCA CATTCTCACG GGATTTCAGTC GTCACTCATG GTGATTTCCTC ACTTGATAAC
 1501 CTTATTTTG ACGAGGGGAA ATTAATAGGT TGTTATTGATG TTGGACGAGT CGGAATCGCA
 1561 GACCGATACC AGGATCTTGC CATCCATGG AACTGCCTCG GTGAGTTTTC TCCTTCATTA
 1621 CAGAAACGGC TTTTTCAAAA ATATGGTATT GATAATCTG ATATGAATAA ATTGCAAGTTT
 1681 CATTGATGCA TCGATGAGTT TTTCTAATCA GAATTGGTTA ATTGGTTGTA ACATTATTCA
 1741 GATTGGGCC CGTTCACGAG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA
 1801 GATCCTTTT TTCTGCGCGT AATCTGCTG TTGCAAAACAA AAAAACCAAC GCTACCGCG
 1861 GTGGTTTGTG TGCGGATCA AGAGCTACCA ACTCTTTTC CGAAGGTAAC TGGCTTCAGC
 1921 AGAGCGCAGA TACCAAAATC TGTTCTCTA GTGTAGCCGT AGTTAGGCCA CCACCTCAAG
 1981 AACTCTGTAG CACCGCTAC ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC
 2041 AGTGGCGATA AGTCGTGTCT TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG
 2101 CAGCGGTCGG GCTGAACGGG GGGTTCTGAC ACACAGCCCA GCTTGGAGCG AACGACCTAC
 2161 ACCGAACCTGA GATACTACA GCGTGAGCTA TGAGAAAGCG CCACGCTTCC CGAAGGGAGA
 2221 AAGGCAGACA GGTATCCGGT AAGCGGCAGG GTCCGAACAG GAGAGCGCAC GAGGGAGCTT
 2281 CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTGCGCACCT CTGACTTGAG
 2341 CGTCGATTTT TGTGATGCTC GTCAAGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACCGCG
 2401 GCCTTTTTAC GGTTCTGGC CTTTTGCTGG CCTTTGCTC ACATGTTCTT TCCTGCGTTA
 2461 TCCCCTGATT CTGTGGATAA CCGTATTACC GCTAGCATGG ATCTCGGGGA CGTCTAACTA
 2521 CTAAGCGAGA GTAGGGAACT GCCAGGCATC AAATAAAACG AAAGGCTCAG TCGGAAGACT
 2581 GGGCTTTGCG TTTTATCTGT TGTTTGTGG TGAACGCTCT CCTGAGTAGG ACAAATCCGC
 2641 CGGGAGCGGA TTTGAACGTT GTGAAGCAAC GGGCCGGAGG GTGGCGGGCA GGACGCCGC
 2701 CATAAAACTGC CAGGCATCAA ACTAAGCAGA AGGCCATC

Fig 16B

24/260

Figure 17A: Cloning Sites of the *E1t*_{LY} Vector pETURB

Int attL

--- tat tac aaa aaa gca ggc ttt gaa aac ctg tat ttt caa gya
gtt gtc ttt cgt ccg aaa ctt ttg gac ata ana gtt cct
 Leu Tyr Lys Lys Ala Gly Phe Glu Asn Leu Tyr Phe Gln, Gly

TEV Protease

NcoI Apa II Sal I BamH I KpnI EcoI
 atc atg gac cta gtc gac tgg atc egg tac cda att cgc ---
 tgg tac ctg gat cag ctg acc tag gcb atg gct taa gcg ---
 Thr Met Asp Leu Val Asp Trp Ile Arg Tyr Arg Ile

EcoI XbaI XbaI EcoI XbaI attL
 Death --- aga att cgc ggc cgc act cga gat acc tag acc cag
 --- tct taa ggc cgg ggc tga gtc cta tag atc tgg gtc

Int
att
ctt tcc gtc aca ttt ---
gaa aga aca tgt ttc ---

25/240

pENTR8 2735 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
339..644	ccdB
673..772	attL2
895..1704	KmR
1809..2382	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
 61 GGGCCCCAAA TAATGATTG ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAAATTGAT
 121 AAGCAATGCT TTTTATAAT GCCAACTTTG TACAAAAAAG CAGGCTTTGA AAACCTGTAT
 181 TTTCAGGAA CCATGGACT AGTCGACTGG ATCCGGTACCC GAATTGCGTT ACTAAAAGCC
 241 AGATAAACAGT ATGCGTATTT GCGCGCTGAT TTTTGCCTGA TAAGAATATA TACTGATATG
 301 TATACCCGAA GTATGTCAA AAGAGGTGTG CTCTAGAAT GCAGTTAAG GTTTACACCT
 361 ATAAAAAGAGA GAGCCGTTAT CGTCTGTTT TGATGTACA GAGTGATATT ATTGACACGC
 421 CCGGGCGACG GATAGTGATC CCCCTGGCCA GTGCGACGTCT GCTGTCAGAT AAAAGTCCTCC
 481 GTGAACTTTA CCCGGTGGTG CATATCGGGG ATGAAAGCTG GCGCATGATG ACCACCGATA
 541 TGGCCAGTGT GCGGGCTCC GTTATCGGGG AAGAAGTGGC TGATCTCAGC CACCGCGAAA
 601 ATGACATCAA AAACGCCATT AACCTGATGT TCTGGGGAAAT ATAGAATTGCG ACCGGCCACT
 661 CGAGATATCT AGACCCAGCT TTCTTGTACA AAGTTGGCAT TATAAGAAAG CATTGCTTAT
 721 CAATTGTTG CAACGAACAG GTCACTATCA GTCAAAATAA AATCATTATT TGCCATCCAG
 781 CTGCAGCTCT GGCCCGTGTG TCAAAATCTC TGATGTTACA TTGCAACAAGA TAAAAATATA
 841 TCATCATGAA CAATAAAACT GTCTGTTAC ATAACAGTA ATACAAGGGG TGTTATGAGC
 901 CATATTCAAC GGGAAACGTC GAGGCCGCGA TAAATTCGA ACATGGATGC TGATTATAT
 961 GGGTATAAAT GGGCTCGCGA TAATGTCGGG CAATCAGGTG CGACAACTCA TCGCTTGTAT
 1021 GGGAAAGCCCG ATGCGCCAGA GTTGTGCTG AAACATGGCA AAGGTAGCGT TGCCAAATGAT
 1081 GTTACAGATG AGATGGTCAG ACTAAACTGG CTGACGGAAAT TTATGCTCT TCCGACCATC
 1141 AACGATTTA TCCGTAATCC TGATGATGCA TGTTACTCA CCACTGCGAT CCCCCGGAAA
 1201 ACAGCATTCC AGGTATTAGA AGAATATCCT GATTCAAGGTG AAAATATTGT TGATGCGCTG
 1261 CGAGTGTCCC TGCGCCGGTT GCATTGATT CCTGTTGTA ATTGTCCTTT TAACAGCGAT
 1321 CGCGTATTC GTCTCGCTCA GGCGCAATCA CGAATGAATA ACGGTTGGT TGATGCGAGT
 1381 GATTTGATG ACGAGCGTAA TGGCTGGCCT GTTGAACAAG TCTGAAAGA AATGCATAAA
 1441 CTTTGCCAT TCTCACCGGA TTCAGTCGTC ACTCATGGTG ATTTCTCACT TGATAACCTT
 1501 ATTTTGACG AGGGGAAATT AATAGGTTGT ATTGATGTTG GACGAGTCGG AATCGCAGAC
 1561 CGATACCAGG ATCTGCCAT CCTATGGAAC TGCCCTCGGTG AGTTTTCTCC TTCATTACAG
 1621 AAAAGGTTT TCTAAAAATA TGGTATTGAT AATCCTGATA TGAATAAATT GCAGTTTCAT
 1681 TTGATGCTCG ATGAGTTTT CTAATCAGAA TTGGTTAATT GGTTGTAACA TTATTGAGAT
 1741 TGGGCCCCGT TCCACTGAGC GTCAGACCCC GTAGAAAAAGA TCAAAAGGATC TTCTTGAGAT
 1801 CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCACCGCT ACCACGGGTG
 1861 GTTGTGTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACTGG CTTFCAGCAGA
 1921 GCGCAGATAC CAAATACTGT TCTCTAGTG TAGCCGTAGT TAGGCCACCA CTTCAAGAAC
 1981 TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCACTGGC TGCTGCCAGT
 2041 GGGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA TAAGGCGCAG
 2101 CGGTCGGGCT GAACGGGGGG TTCGTCACCA CAGCCCAGCT TGGAGCGAAC GACCTACACC
 2161 GAAGTGGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA AGGGAGAAAG
 2221 GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCAGGAG GGAGCTTCCA
 2281 GGGGGAAACG CCTGGTATCT TTATAGTCTC GTCCGGTTTC GCCACCTCTG ACTTGAGCGT
 2341 CGATTTTGTG ATGCTCGTC AGGGGGCGG AGCCTATGGA AAAACGCCAG CAACGCCGCC
 2401 TTTTTACGGT TCCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTCC TGCGTTATCC
 2461 CCTGATTCTG TGGATAACCG TATTACCGCT AGCATGGATC TCGGGGACGT CTAACACTA
 2521 AGCGAGAGTA GGGAACTGCC AGGCATCAAA TAAAACGAAA GGCTCAGTCG GAAGACTGG
 2581 CCTTTGCTTT TATCTGTTGT TTGTCGGTGA ACGCTCTCCT GAGTAGGACA AATCCGGGG
 2641 GAGCGGATTT GAACGTTGTG AAGCAACGGC CGGGAGGGTG CGGGGCAGGA CGCCCGCCAT
 2701 AAACTGCCAG GCATCAAAC AAGCAGAAGG CCATC

FIGURE 17B

26/240

Figure 18A: Cloning sites of the *ENTY* Vector pENTR9

Int attL1

ttg tac aaa aaa gca ggc ttt gaa aac ctg tat ttt caa gga
 paa aag ttt ttc cgt ccg aaa ctt ttg gac ata aaa gtt cct
 Leu Tyr Lys Lys Ala Gly Phe Glu Asn Leu Tyr Phe Gln, Gly
 TEV protease

NdeI BglII SalI BamHI KpnI EcoRI
 cat atg afa tot gtc gac tgg atc cgg tac cod att cgc ---
 gta tac tct aga cag cgg acc tag gac atg get taa geg ---
 His Met Arg Ser Val Asp Trp Ile Arg Tyr Arg Ile

EcoRI NheI XbaI EcoRI NheI attL2

DeoR --- aga att cgc ggc ege act cga gat atc tag acc cag
 --- tct taa gcg eeg ggg tga gct cta tag atc tgg gtc

Int

27/240

pENTR9 2735 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
339..644	ccdB
673..772	attL2
895..1704	KmR
1809..2382	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
 61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT
 121 AAGCAATGCT TTTTATAAT GCCAACCTTG TACAAAAAAAG CAGGCTTGA AAACCTGTAT
 181 TTCAAGGAC ATATGAGATC TGTCGACTGG ATCCGGTACCC GAATTCGCTT ACTAAAAGCC
 241 AGATAACAGT ATCGTGTATTG CGCGCCTGAT TTTTGCCTGA TAAGAATATA TACTGATATG
 301 TATACCCGAA GTATGTCAAAGAGGGTGTG CTCTCTAGAAT GCAGTTAAG GTTTACACCT
 361 ATAAAAGAGA GAGCGGTAT CGTCTGTTG TGGATGTACA GAGTGATATT ATTGACACGC
 421 CCGGGCGACG GATAGTGAAT CCCCTGGCCA GTGGAGGCTC GCTGTAGAT AAAGTCTCCC
 481 GTGAACTTTA CCCGGTGGTG CATATCGGGG ATGAAAGCTG GCGCATGATG ACCACCGATA
 541 TGCCAGTGTG GCGGTCTCC GTTATCGGGG AAGAAGTGGC TGATCTCAGC CACCGCGAAA
 601 ATGACATCAA AAACGCCATT AACCTGATGT TCTGGGAAAT ATAGAATTG CGGCCGCACT
 661 CGAGATATCT AGACCCAGT TTCTTGATCA AGATGGCAT TATAAGAAAG CATTGCTTAT
 721 CAATTGTTG CAACGAACAG GTCACTATCA GTCAAAATAA AATCATTATT TGCCATCCAG
 781 CTGCAGCTCT GGCCCGTGT TCAAAATCTC TGATGTTACA TTGCACAAGA TAAAATATA
 841 TCATCATGAA CAATAAAACT GTCTGCTTAC ATAAACAGTA ATACAAGGGG TGTTATGAGC
 901 CATATTCAAC GGGAAACGTC GAGGCCGCGA TTAAATTCCA ACATGGATGC TGATTTATAT
 961 GGGTATAAAT GGGCTCGCGA TAATGTCGGG CAATCAGGTG CGACAATCTA TCGCTTGTAT
 1021 GGGAAAGCCCG ATGCGCCAGA GTTGTGTTCTG AAACATGGCA AAGGTAGCGT TGCCAAATGAT
 1081 GTTACAGATG AGATGGTCAG ACTAAACTGG CTGACGGAAT TTATGCTCT TCCGACCAC
 1141 AAGCATTITA TCCGTACTCC TGATGATGCA TGGTTACTCA CCACTGCGAT CCCCCGAAAA
 1201 ACAGCATTCC AGGTATTAGA AGAATATCCT GATTCAAGGTG AAAATATTGT TGATGCGCTG
 1261 GCAGTGTCCC TGCGCCGGTT GCATTGATT CCTGTTGTA ATTGCTCTT TAACAGCGAT
 1321 CGCGTATTTC GTCTCGCTCA GGCGCAATCA CGAATGAATA ACGGTTGGT TGATGCGAGT
 1381 GATTTGATG ACGAGCGTAA TGGCTGGCT GTGAAACAAG TCTGAAAGA AATGCATAAA
 1441 CTTTGCCAT TCTCACCGGA TTCAGTCGTC ACTCATGGTG ATTTCTCACT TGATAACCTT
 1501 ATTTTGACG AGGGGAAATT AATAGGTTGT ATTGATGTTG GACGAGTCGG AATCGCAGAC
 1561 CGATACCAGG ATCTTGCCAT CCTATGGAAC TGCCTCGGTG AGTTTCTCC TTCATTACAG
 1621 AAACGGCTTT TTCAAAATAA TGGTATTGAT AACCTGTATA TGAATAAATT GCAGTTTCAT
 1681 TTGATGCTCG ATGAGTTTT CTAATCAGAA TTGGTTAATT GGTTGTAACA TTATTCAGAT
 1741 TGGGCCCGT TCCACTGAGC GTCAGACCCCC GTAGAAAAAGA TCAAAAGGATC TTCTTGAGAT
 1801 CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCACCGCT ACCAGCGGTG
 1861 GTTGTGTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACTGG CTTCAGCAGA
 1921 GCGCAGATAC CAAATACGTG TCTTCTAGTG TAGCCGTAGT TAGGCCACCA TTCAAGAAC
 1981 TCTGTAGCAC CGCCTACATA CTCCTGCTCTG CTAATCCTGT TACCACTGGC TGCTGCCAGT
 2041 GGGGATAAGT CGTGTCTTAC CGGGTGGAC TCAAGACGAT AGTTACCGGA TAAGGCGCAG
 2101 CGGTGGGCT GAACGGGGGG TTCTGCAACAC GACCCAGCT TGGAGCGAAC GACCTACACC
 2161 GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAGCGCCA CGCTTCCGA AGGGAGAAAAG
 2221 GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGGAG AGCGCACGAG GGAGCTTCCA
 2281 GGGGGAAACG CCTGGTATCT TTATGCTCCT GTGGGGTTTC GCCACCTCTG ACTTGAGCGT
 2341 CGATTTTGTG ATGCTCGTC AGGGGGCGG AGCTTATGGA AAAACGCCAG CAACGCCGG
 2401 TTTTACGGT TCCTGGCTT TTGCTGGCT TTGCTCACA TGTTCTTCC TGCGTTATCC
 2461 CCTGATTCTG TGGATAACCG TATTAACCGCT AGCATGGATC TCGGGGACGT CTAACACTA
 2521 AGCGAGAGTA GGGAACTGCC AGGCATCAAATAAAACGAAA GGCTCAGTCG GAAGACTGGG
 2581 CCTTCGTTT TATCTGTTGT TTGCTGGTGA ACGCTCTCCT GAGTAGGACA AATCCGCCGG
 2641 GAGCGGATTT GAACGTTGTG AACCAACGGC CCGGAGGGTG GCGGGCAGGA CGCCCGCCAT
 2701 AAACTGCCAG GCATCAAAC AAGCAGAAGG CCATC

FIGURE 18B

28/260

Figure 19A: Cloning sites of the ENTRY Vector pENTR10

Int attL1 S.D. - - 12 NdeI

--- atg tac aaa aaa gca ggc ttc gaa cta agg aaa tac tta cat
--- agg atg atc ttt cgt ccc ang ctt gat tcc ttt atg aat gta

Leu Tyr Lys Lys Ala Gly Phe Glu Leu Arg Lys Tyr Leu His

Kpn Xba Sd Bam Kpn EcoRI

atg gga [acc] aat tca gtc gac tgg atc cgg tac [cg] aat cgc ---
 tac cct tgg tta agt cag ctc acc tag [g] atg get taa [g] ggc ---
 Met Gly Thr Asn Ser Val Arg Trp Ile Arg Tyr Arg Ile

EcoRI Not Xba EcoRI Xba attL2

Death --- aat att cgc [ggc cgc act cga gat] atc tag [acc cag]
 (ccdB) --- tct taa [g] cgc ccc tga gct cta tag atc [tgg gtc]

Int

--- tct tgg ggc acc tag ---
gaa aga aca cgt ttc ---

29/240

pENTR10 2738 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
342..647	ccdB
676..775	attL2
898..1707	KmR
1812..2385	ori

1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
 61 GGGCCCAA TAATGATT TTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT
 121 AAGCAATGCT TTTTATAAT GCCAACCTTG TACAAAAAAG CAGGCTTCGA ACTAAGGAAA
 181 TACCTACATA TGGGAACCAA TTCAGTCGAC TGAGATCCGGT ACCGAATTGCT CTTACTAAAA
 241 GCCAGATAAC AGTATGCGTA TTTGCGCGCT GATTTTTGCG GTATAAGAAT ATATACTGAT
 301 ATGTATACCC GAAGTATGTC AAAAAGAGGT GTGCTTCTAG AATGCAGTTT AAGGTTTACA
 361 CCTATAAAAG AGAGAGCCGT TATCGTCTGT TTGTGGATGT ACAGAGTGAT ATTATTGACA
 421 CGCCCAGGGCG ACAGGATGGTG ATCCCCCTGG CCAGTGCACG TCTGCTGTCA GATAAAGTCT
 481 CCCGTGAACCT TTACCCGGTG GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCCACCG
 541 ATATGGCCAG TGTGCCGGTC TCCGTTATCG GGGAAAGAAG GGCTGATCTC AGCCACCGCG
 601 AAAATGACAT CAAAAACGCC ATTAACCTGA TGGTCTGGGG AATATAGAAT TCGCGGCCCGC
 661 ACTCGAGATA TCTAGACCCA GCTTTCTTGT ACAAAAGTGG CAFTATAAGA AAGCATTGCT
 721 TATCAATTG TTGCAACGAA CAGGTCACTA TCAGTCAAA TAAATCATT ATTTGCCATC
 781 CAGCTGCAGC TCTGGCCCGT GTCTCAAAAT CTCTGATGTT ACATTGCAACA AGATAAAAAT
 841 ATATCATCAT GAACAATAAAA ACTGCTCTGCT TACATAAAACA GTAATACAAG GGGTCTTATG
 901 AGCCATATTC AACGGGAAAC GTCGAGGCCG CGATTAAAATT CCAACATGGA TGCTGATTAA
 961 TATGGGTATA AATGGGCTCG CGATAATGTC GGGCAATCAG GTGCGACAAT CTATCGCTTG
 1021 TATGGGAAGC CGGATGCGCC AGAGTTGTTT CTGAAACATG GCAAAGGTAG CGTTGCCAAT
 1081 GATGTTACAG ATGAGATGGT CAGACTAAAC TGCTGACGG AATTTATGCC TCTTCCGACC
 1141 ATCAAGCATT TTATCCGTAC TCCTGATGAT GCATGGTTAC TCACCACTGC GATCCCCGA
 1201 AAAACAGCAT TCCAGGTATT AGAAGAAATAT CCTGATTTCAG GTGAAAATAT TGTTGATGCG
 1261 CTGGCAGTGT TCCCTGCGCC GTTGCAATTG ATTCTGTTT GAAATTGTC TTTTAACAGC
 1321 GATCGCGTAT TTGCGTCTGC TCAGGGCGAA TCACGAATGA ATAACGGTTT GGTTGATGCG
 1381 AGTGTATTTG ATGACGAGCG TAATGGCTGG CCTGTTGAAC AAGTCTGAA AGAAATGCAT
 1441 AAACCTTTGC CATTCTCACCG GATTTCAGTC GTCACTCATG GTGATTCTC ACTTGATAAC
 1501 CTTATTTTG ACGAGGGGAA ATTAATAGGT TGTATTGATG TTGGACGAGT CGGAATCGCA
 1561 GACCGATACC AGGATCTTGC CATCCTATGG AACTGCCTCG GTGAGTTTTC TCCTTCATTA
 1621 CAGAAACGGC TTTTCAAAA ATATGGTATT GATAATCTG ATATGAATAA ATTGCAAGTTT
 1681 CATTGAGTT TTTCTAATCA GAATTGGTTA ATTGGTTGTA ACATTATTCA
 1741 GATTGGGCC CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTGGA
 1801 GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA AAAAACCAAC GCTACCAGCG
 1861 GTGGTTGTT TGCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC
 1921 AGAGCGCAGA TACCAAATAC TGTTCTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAG
 1981 AACTCTGTAG CACCGCCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAAGT GGCTGCTGCC
 2041 AGTGGCGATA AGTCGTGCT TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG
 2101 CAGCGGTCGG GCTGAACGGG GGGTTCTGTC ACACAGCCCA GCTGGAGCG AACGACCTAC
 2161 ACCGAACCTGA GATACCTACA GCGTAGAGCTA TGAGAAAGCG CCACGCTTCC CGAAGGGAGA
 2221 AAGGGCGACA GGTATCCGGT AAGCAGCAGG GTCGGGACAG GAGAGCGCAC GAGGGAGCTT
 2281 CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTGCCCCACT CTGACTTGG
 2341 CGTCGATTT TGTGATGCTC GTCAAGGGGG CGGAGCCTAT GGAAAACGC CAGCAACGCG
 2401 GCCTTTTTAC GGTTCCCTGGC CTTTTGCTGG CCTTTTGTCTC ACATGTTCTT TCCTGCGTTA
 2461 TCCCCCTGATT CTGTGGATAA CCGTATTACG GCTAGCATGG ATCTCGGGGA CGTCTAACTA
 2521 CTAAGCGAGA GTAGGGAACT GCCAGGCATC GAATAAAACG AAAGGCTCAG TCAGGAAGACT
 2581 GGGCCTTCG TTTTATCTGT TTGTTGTCGG TGAACGCTCT CCTGAGTAGG ACAAATCCGC
 2641 CGGGAGCGGA TTGAACTGTT GTGAAGCAAC GGCCCCGGAGG GTGGCGGGCA GGACGCCCGC
 2701 CATAAAACTGC CAGGCATCAA ACTAAGCAGA AGGCCATC

FIGURE 19B

30/240

Figure 20A: Cloning Sites of the Entry Vector pENTR111

Int	attL1	S.D.	Kozak	XmnI	S.D.
TTG	TAC AAA AAA GCA GGC	TTC GAA GGA GAT AGA ACC	AAT TCT CTA AGG AAA TAC		
AAC ATG TTT	TTT CGT CCG AAG CTT CCT CTA TCT TGG	TTA AGA GAT TCC TTT ATG			
Leu Tyr Lys Lys Ala Gly Phe Glu Gly Asp Arg Thr Asn Ser Leu Arg Lys Tyr					

Kozak	NcoI	Sall	BamHI	KpnI	EcoRI	EcoRI	NotI
TTA	ACC ATG	GTC GAC	TGG ATC	CGG TAC	CGA ATT C	--ccdB	--G
AAT	TGG TAC	CAG CTG	ACC TAG	GCC ATG	GCT TAA G	C	AAT TCG dGG CCG
	↓	↓	↓	↓	↓	TTA AGC	GCC GGC
						↓	↓
Leu Thr Met Val Asp Trp Ile Arg Tyr Arg Ile						Asn Ser Arg Pro	

XbaI	EcoRV	XbaI	Int	attL2
CAC	TCG AGA	TAT	CTA GAC	CCA GCT TTC TTG TAC AAA G
GTG	AGC TCT	ATA GAT	CTG GGT CGA AAG AAC ATG	TTT C
↓	↓	↓		
His Ser Arg Tyr Leu Asp Pro Ala Phe Leu Tyr Lys				

31/240

pENTR11 2744 bp (rotated to position 2578)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
348..653	ccdB
683..781	attL2
904..1713	KmR
1818..2391	ori

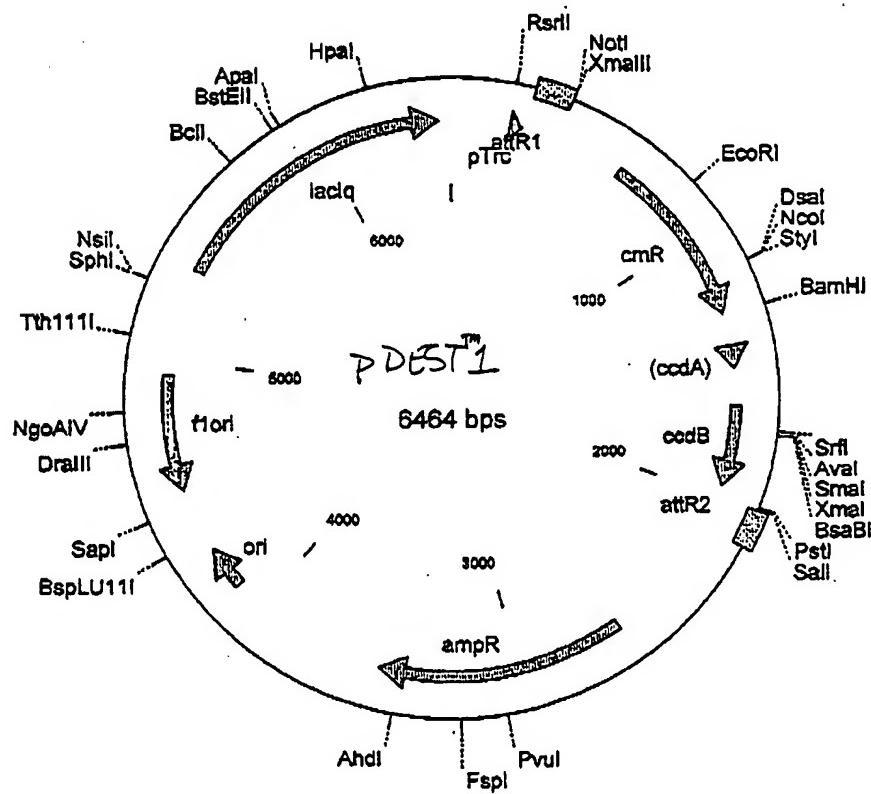
1 CTGACGGATG GCCTTTTGC GTTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
 61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAATTGAT
 121 AAGCAATGCT TTTTATAAT GCCAACTTTG TACAAAAAAG CAGGCTTCGA AGGAGATAGA
 181 ACCAATTCTC TAAGGAATA CTTAACCATG GTCGACTGGA TCCGGTACCG AATTGCTTA
 241 CTAAAAGCCA GATAACAGTA TGCGTATTTG CGCGCTGATT TTTGCGGTAT AAGAATATAT
 301 ACTGATATGT ATACCCGAAG TATGTCAAAA AGAGGTGTGC TTCTAGAATG CAGTTAAGG
 361 TTACACCTA TAAAAGAGAG AGCCGTTATC GTCTGTTGT GGATGTACAG AGTGTATTA
 421 TTGACACGCC CGGGCGACGG ATAGTGTATCC CCCTGGCCAG TGCACGCTG CTGTCAGATA
 481 AAGTCTCCCG TGAACCTTAC CCGGTGGTGC ATATCGGGG TGAAAGCTGG CGCATGATGA
 541 CCACCGATAT GGCCAGTGTG CCGGTCTCCG TTATCGGGG AGAAGTGGCT GATCTCAGCC
 601 ACCCGGAAAA TGACATCAA AACGCCATTA ACCTGTATGTT CTGGGGATA TAGAATTGCG
 661 GGGCGCACTC GAGATATCTA GACCCAGCTT TCTTGTACAA AGTTGGCATT ATAAGAAAGC
 721 ATTGCTTATC AATTGTTGC AACGAACAGG TCACTATCAG TCAAAATAAA ATCATTATTT
 781 GCCATCCAGC TGCACTCTG GCCCCTGTCT CAAAATCTCT GATGTTACAT TGCAACAAGAT
 841 AAAAATATAT CATCATGAA AATAAAACTG TCTGCTTACA TAAACAGTAA TACAAGGGT
 901 GTTATGAGCC ATATTCAACG GGAAACGTCC AGGCCCGGAT TAAATTCAA CATGGATGCT
 961 GATTATATG GGTATAATG GGCTCGCGAT ATGTCGGGC AATCAGGTGC GACAATCTAT
 1021 CGCTGTATG GGAAGCCCGA TGCGCCAGAG TTGTTCTGA AACATGGCAA AGGTAGCGTT
 1081 GCCAATGATG TTACAGATGA GATGGTCAGA CTAACACTGGC TGACGGAATT TATGCTCTT
 1141 CCGACCATCA AGCATTATAT CCGTACTCCT GATGATGCGAT GGTTACTCAC CACTGCGATC
 1201 CCCGGAAAAA CAGCAATTCA GGTATTAGAA GAATATCTG ATTCAGGTGA AAATATTGTT
 1261 GATGCGCTGG CAGTGTTCCT GCGCCGGTGT CATTGCGATTC CTGTTGTAA TTGTCCTTTT
 1321 AACAGCGATC GCGTATTTCG TCTCGCTCAG GCGCAATCAC GAATGAATAA CGGTTGGTT
 1381 GATGCGAGTG ATTTGATGA CGAGCGTAAT GGCTGGCTG TTGAAACAAGT CTGGAAAGAA
 1441 ATGCATAAAC TTTGCCATT CTCACCGGAT TCAGTCGTCA CTCATGGTGA TTTCTCACTT
 1501 GATAACCTA TTTTGACGA GGGGAAATTA ATAGGTTGTA TTGATGTTGG ACGAGTCGGA
 1561 ATCGCAGACC GATACCAGGA TCTTGCCATC CTATGGAACT GCCTCGGTGA GTTCTCTCCT
 1621 TCATTACAGA AACGGTTTT TCAAAATAT GGTATTGATA ATCCGTATAT GAATAAATTG
 1681 CAGTTTCATT TGATGCTCGA TGAGTTTTT TAATCAGAAT TGGTTAATTG GTTGTACAT
 1741 TATTTCAGATT GGGCCCCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGGATCT
 1801 TCTTGAGATC CTTTTTTCTC GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA
 1861 CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC TTTTCCGAA GGTAACGGC
 1921 TTCAGCAGAG CGCAGATACC AAATACTGTT CTCCTAGTGT AGCCGTAGTT AGGCCAC
 1981 TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT
 2041 GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGTTGGACT CAAGACGATA GTTACCGGAT
 2101 AAGGCGCAGC GGTCGGGTG AACGGGGGT TCGTGCACAC AGCCCAGCTT GGAGCGAAC
 2161 ACTTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCAC GCTTCCGAA
 2221 GGGAGAAAGG CGGACAGGTG TCCGCTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG
 2281 GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCTG TCGGGTTTCG CCACCTCTGA
 2341 CTTGAGCGTC GATTTTTGTT ATGCTCGTCA GGGGGGGCGGA GCCTATGAA AAACGCCAGC
 2401 AACGGGGCCTT TTTTACGGTT CCTGGCCCTT TGCTGGCCTT TTGCTCACAT GTTCTTCCCT
 2461 GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCTA GCATGGATCT CGGGGACGTC
 2521 TAACTACTAA GCGAGAGTAG GGAACCTGCCA GGCATCAAAT AAAACGAAAG GCTCAGTCGG
 2581 AAGACTGGGC CTTTCGTTTT ATCTGTTGTT TGTCGGTGAA CGCTCTCTG AGTAGGACAA
 2641 ATCCGCCGGG AGCGGATTG AACGTTGTGA AGCAACGGCC CGGAGGGTGG CGGGCAGGAC
 2701 GCCCGCCATA AACTGCCAGG CATCAAACTA AGCAGAAGGC CATC

FIGURE 20B

32/240

Figure 2: A:pDESTTM Native Protein Expression in E. coli

1 atgagctgt **-35 Trc promoter** gacaattaaat catccggctc **-10** **rRNA**
 tactcgacaa ctgttaatta gtaggccag catattcac accttaaacac tccgcatttg
 61 aatttcacac agaaaacaga caggatagg atccaaaggtt **Tat attR1** tttagaaaaa/actgttggaa
 tttaagtgtg tcctttgtct gtcataatcc tagtgttcaa acatgttttc pggatcgtc



33/240

pDEST1 6464 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
216..257	Trc promoter
397..273	attR1
647..1306	CmR
1426..1510	inactivated ccdA
1648..1953	ccdB
1994..2118	attR2
2598..3503	ampR
4104..4264	ori
4504..4941	flori (f1 intergenic region)
5340..6420	lacIq

1 GTTTGACAGC TTATCATCGA CTGCACGGTG CACCAATGCT TCTGGCGTCA GGCAGCCATC
 61 GGAAGCTGTG GTATGGCTGT GCAGGTCGTA AATCACTGCA TAATTCGTGT CGCTCAAGGC
 121 GCACTCCCGT TCTGGATAAT GTTTTTGCG CCGACATCAT AACGGTTCTG GCAAATATTG
 181 TGAAATGAGC TGTGACAAT TAATCATCCG GTCCGTATAA TCTGTGGAAT TGTGAGCGGG
 241 ATAACAATTT CATCGCGAGG TACCAAGCTA TCACAAGTTT GTACAAAAAA GCTGAACGAG
 301 AAACGTAAAA TGATATAAAT ATCAATATAT TAAATTAGAT TTTGCATAAA AAACAGACTA
 361 CATAATACTG TAAAACACAA CATATCCAGT CACTATGGCG GCGCCTAAGT TGGCAGCATT
 421 ACCCGACGCA CTTTGCGCC AATAAAATACC TGTGACGGAA GATCACTTCG CAGAATAAAAT
 481 AAATCCTGGT GTCCCTGTTC ATACCGGGAA GCCCTGGGCC AACTTTGGC GAAAATGAGA
 541 CGTTGATCGG CACGTAAGAG GTTCCAACCT TCACCATATAA GAAATAAGAT CACTACCGG
 601 CGTATTTTTT GAGTTATCGA GATTTTCAAG AGCTAAGGAA GCTAAAATGG AGAAAAAAAT
 661 CACTGGATAT ACCACCGTTG ATATATCCCA ATGGCATCGT AAAAGAACATT TTGAGGCATT
 721 TCAGTCAGTT GCTCAATGTA CCTATAAACCA GACCGTTCAAG CTGGATATTAA CGGCCTTTTT
 781 AAAGACCGTA AAGAAAAATA AGCACAAGTT TTATCCGGCC TTTATTACAA TTCTTGGCCG
 841 CCTGATGAAT GCTCATCCGG AATTCCGTAT GGCAATGAAA GACGGTGAGC TGGTGTATAG
 901 GGATAGTGTT CACCCCTGTT ACACCGTTT CCATGAGCAA ACTGAAACGT TTTCATCGCT
 961 CTGGAGTGAA TACCAACGACC ATTTCCGGCA GTTTCTACAC ATATATTGCG AAGATGTGGC
 1021 GTGTTACGGT GAAAACCTGG CCTATTTCCC TAAAGGGTTT ATTGAGAATA TGTTTTTCGT
 1081 CTCAGCCAAT CCCTGGGTGA GTTTCACCAAG TTTGATTTA AACGTGGCCA ATATGGACAA
 1141 CTTCTTCGCC CCCGTTTCA CCATGGGCAA ATATTATACG CAAGGGGACA AGGTGCTGAT
 1201 GCCGCTGGCG ATTCAAGGTT ATCATGCCGT CTGTGATGGC TTCCATGTGCG GCAGAAATGCT
 1261 TAATGAATTAA CAACAGTACT GCGATGAGTG GCAGGGCGGG GCGTAAACGCG GTGGATCCGG
 1321 CTTACTAAAA GCCAGATAAC AGTATGCGTA TTTGCGCGCT GATTTTGCG GTATAAGAAT
 1381 ATATACTGAT ATGTATACCC GAAGTATGTC AAAAAGAGGT GTGCTATGAA GCAGCGTATT
 1441 ACAGTGACAG TTGACAGCGA CAGCTATCAG TTGCTCAAGG CATATATGAT GTCAATATCT
 1501 CCGGTCTGGT AAGCACAACC ATGCAGAATG AAGCCCGTCG TCTGCGTGC GAACGCTGGA
 1561 AAGCGAAAAA TCAGGAAGGG ATGGCTGAGG TCGCCCGGTT TATTGAAATG AACGGCTCTT
 1621 TTGCTGACGA GAAACAGGGAC TGGTGAATG CAGTTTAAGG TTTACACCTA TAAAAGAGAG
 1681 AGCCGTTATC GTCTGTTTGT GGATGTACAG AGTGTATTTA TTGACACGCC CGGGCGACGG
 1741 ATGGTGTATCC CCCTGGGCCAG TGACAGCTCTG CTGTCAGATA AAGTCTCCCG TGAACCTTAC
 1801 CCGGTTGTC ATATCGGGGA TGAAAGCTGG CGCATGATGA CCACCGATAT GGCCAGTGTG
 1861 CCGGCTCTCCG TTATCGGGGA AGAAGTGGCT GATCTCAGCC ACCGCGAAAAA TGACATCAAA
 1921 AACGCCATTA ACCTGATGTT CTGGGGAAATA TAAATGTCAG GCTCCCTTAT ACACAGCCAG
 1981 TCTGCAAGTC GACCATACTG ACTGGATATG TTGTTGTTTA CAGTATTATG TAGTCTGTTT
 2041 TTTATGCAAA ATCTAATTAA ATATATTGAT ATTTATATCA TTTTACGTTT CTGTTCACTG
 2101 TTCTTGTAC AAAGTGGTGA TAGCTTGGCT GTTTTGGCGG ATGAGAGAAAG ATTTTCAGCC
 2161 TGATACAGAT TAAATCAGAA CGCAGAACGCG GTCTGATATAA ACAGAAATTG CCTGGCGGCC
 2221 GTAGCGCGGT GGTCCCCACCT GACCCCATGC CGAACCTCAGA AGTGAACGCC CGTAGCGCCCG
 2281 ATGGTAGTGT GGGGTCTCCC CATGCGAGAG TAGGGAACCTG CCAGGCATCA AAAAAAACGA
 2341 AAGGCTCAGT CGAAAGACTG GGCCTTCTCGT TTATCTGTT GTTTGTCGGT GAACGCTCTC
 2401 CTGAGTAGGA CAAATCCGCC GGGAGCGGAT TTGAACGTTG CGAACGAAACG GCCCCGGAGGG
 2461 TGGCGGGCGAG GACGCCCGCC ATAAACTGCC AGGCATCAA TTAAGCAGAA GGCCATCTG
 2521 ACGGATGGCC TTTTGCGTT TCTACAAACT CTTTTGTTT ATTGTTCTAA ATACATTCAA-

FIGURE 21B

34/740

2581 ATATGTATCC GCTCATGAGA CAATAACCCCT GATAAAATGCT TCAATAATAT TGAAAAAGGA
 2641 AGAGTATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC CTTTTTGCG GCATTTGCC
 2701 TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA AGATGCTGAA GATCAGTTGG
 2761 GTGCACGAGT GGGTTACATC GAACCTGGATC TCAACAGCGG TAAGATCCTT GAGAGTTTC
 2821 GCCCGAAGA ACGTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT GGCGCGGTAT
 2881 TATCCCCTGT TGACGCCGG CAAGAGAAC TCGGTCGCCG CATAACTAT TCTCAGAAC
 2941 ACTTGGTTGA GTACTCACCA GTCACAGAAA AGCATCTTAC GGATGGCATG ACAGTAAGAG
 3001 AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC GGCCAACTTA CTTCTGACAA
 3061 CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTGCACAA CATGGGGAT CATGTAACTC
 3121 GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG CGTGACACCA
 3181 CGATGCCTAC AGCAATGGCA ACAACGTTGC GCAAACATTAACTGGCGAA CTACTTACTC
 3241 TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA GGACCACCTC
 3301 TGCGCTCGGC CCTTCCGGCT GGCTGGTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG
 3361 GGTCTCGCGG TATCATTGCA GCACCTGGGC CAGATGGTAA GCCCTCCCGT ATCGTAGTTA
 3421 TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA TAGACAGATC GCTGAGATAG
 3481 GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT TTACTCATAT ATACTTTAGA
 3541 TTGATTTAAA ACTTCATTT TAATTTAAA GGATCTAGGT AAAGATCCTT TTTGATAATC
 3601 TCATGACCAA ATCCCTTAA CGTGAGTTT CGTCCACTG AGCGTCAGAC CCCGTAGAAA
 3661 AGATCAAAGG ATCTTCTTGA GATCCTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA
 3721 AAAAACACC GCTACCAGCG GTGGTTGTT TGCGGATCA AGAGCTACCA ACTCTTTTC
 3781 CGAAGGTAAC TGGCTTCAGC AGAGCGAGA TACCAAATAC TGTCTTCTA GTGTAGCCGT
 3841 AGTTAGGCCA CCACTTCAG AACTCTGTAG CACCCCTAC ATACCTCGCT CTGCTAATCC
 3901 TGTTACCACTG GGCTGCTGCC AGTGGCGATA AGTCGTGTCT TACCGGGTTG GACTCAAGAC
 3961 GATAGTTACC GGATAAGGCC CAGCGGTCCG GCTGAACGGG GGGTTCGTGC ACACAGCCCA
 4021 GCTTGGAGCG AACGACCTAC ACCGAACCTGA GATACCTACA GCGTGAGCTA TGAGAAAGCG
 4081 CCACGCTTCC CGAAGGGAGA AAGGCGAGA GGTATCCGGT AAGCGGCAGG GTGGAACAG
 4141 GAGAGCGCAC GAGGGAGCTT CCAGGGGAA ACAGCCTGGTA TCTTATAGT CCTGTCGGGT
 4201 TTCGCCACCT CTGACTTGAG CGTCGATTTT TGTGATGTC GTCAGGGGG CGGAGCCTAT
 4261 GGAAAACGC CAGCAACGCC GCCTTTTAC. GGTTCTGGC CTTTTGCTGG CCTTTGCTC
 4321 ACATGTTCTT TCCCTGCGTTA TCCCCGTATT CTGTTGATAA CCGTATTAC GCCTTGTAGT
 4381 GAGCTGATAC CGCTCGCCGC AGCGAACGA CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG
 4441 CGGAAGAGCG CCTGATGCGG TATTTCTCC TTACGCATCT GTGCGGTATT TCACACCGCA
 4501 TAATTTGTT AAAATCGCG TTAAATTTT GTAAATCAG CTCATTTTT AACCAATAGG
 4561 CCGAAATCGG CAAAATCCCT TATAATCAA AAGAATAGAC CGAGATAGGG TTGAGTGTG
 4621 TTCCAGTTG GAACAAGAGT CCACTATTAA AGAACGTGGA CTCCAACGTC AAAGGGCGAA
 4681 AAACCGTCTA TCAGGGCGAT GGCCCACAC GTGAACCATC ACCCTAATCA AGTTTTTTGG
 4741 GGTGAGGTG CCGTAAAGCA CTAAATCGGA ACCCTAAAGG GAGCCCCCGA TTTAGAGCTT
 4801 GACGGGAAA GCCGGCGAAC GTGGCGAGA AGGAAGGGAA GAAAGCGAAA GGAGCGGGCG
 4861 CTAGGGCGCT GGCAAGTGTG GCGGTACCGC TGCGCGTAAC CACCAACCCC GCCGCGCTTA
 4921 ATGCGCCGCT ACAGGGCGCG TCCATTGCGC ATTCAAGGTG CTATGGTGC CTCTCAGTAC
 4981 AATCTGCTCT GATGCGCGAT AGTTAACCCA GTACAGTCA CGTAGCGATA TCGGAGTGT
 5041 TACACTCCGC TATCGCTACG TGACTGGTC ATGGCTGGC CCGACACCC GCCAACACCC
 5101 GCTGACGCGC CCTGACGGGC TTGCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC
 5161 GTCTCCGGGA GCTGCATGT TCAGAGTTT TCACCGTCAT CACCGAAACG CGCGAGGCAG
 5221 CAGATCAATT CGCGCGCGAA GCGAACCGG CATGCATTAA CGTTGACACC ATCGAATGGT
 5281 GCAAAACCTT TCGCGGTATG GCATGATAGC GCGCGAACAGA GAGTCAAATTC AGGGTGGTGA
 5341 ATGTGAAACC AGTAACGTTA TACGATGTCG CAGAGTATGC CGGTGTCTCT TATCAGACCG
 5401 TTTCCCGCGT GGTGAACCAAG GCCAGCCACG TTTCTGCGAA AACCGGGGAA AAAGTGAAG
 5461 CGCGGATGGC GGAGCTGAAT TACATCCCCA ACCCGTGGC ACAACAACTG GCGGGCAAAC
 5521 AGTCGTTGCT GATTGGCGTT GCCACCTCCA GTCTGGCCCT GCACCGCGCC TCGCAAATTG
 5581 TCGCGCGAT TAAATCTCGC GCCGATCAAC TGGGTGCCAG CGTGGTGGTG TCGATGGTAG
 5641 AACGAAGCGG CGTCGAAGCC TGTAAAGCGG CGGTGCACAA TCTTCTCGCG CAACCGCGTCA
 5701 GTGGGCTGAT CATTAACTAT CCGCTGGATG ACCAGGATGC CATTGCTGTG GAAGCTGCCT
 5761 GCACTAATGT TCCGGCGTTA TTTCTGATG TCTCTGACCA GACACCCATC AACAGTATTA
 5821 TTTTCTCCCA TGAAGACGGT ACGCGACTGG GCGTGGAGCA TCTGGTCGCA TTGGGTACCC
 5881 AGCAAATCGC GCTGTTAGCG GGCCCATTA GTTCTGCTC GCGCGCTCT CGTCTGGCTG
 5941 GCTGGCATAA ATATCTCACT CGCAATCAA TTCAGCCGAT AGCGGAACGG GAAGGGACT
 6001 GGAGTGCCT GTCCGGTTTT CAACAAACCA TGCAAATGCT GAATGAGGGC ATCGTTCCCA-

FIGURE 21C

35/240

6061 CTGCGATGCT GGTTGCCAAC GATCAGATGG CGCTGGGCGC AATGCGCGCC ATTACCGAGT
6121 CCGGGCTGCG CGTTGGTGC GATATCTCGG TAGTGGGATA CGACGATACC GAAGACAGCT
6181 CATGTTATAT CCCGCCGTTA ACCACCATCA AACAGGATT TCGCCTGCTG GGGCAAACCA
6241 GCGTGGACCG CTTGCTGCAA CTCTCTCAGG GCCAGGCGGT GAAGGGCAAT CAGCTGTTGC
6301 CCGTCTCACT GGTGAAAAGA AAAACCAACCC TGGCACCCAA TACGCAAACCC GCCTCTCCCC
6361 GCGCGTTGGC CGATTCA TTATGCAGCTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC
6421 AGTGAGCGCA ACGCAATTAA TGTGAGTTAG CGCGAATTGA TCTG

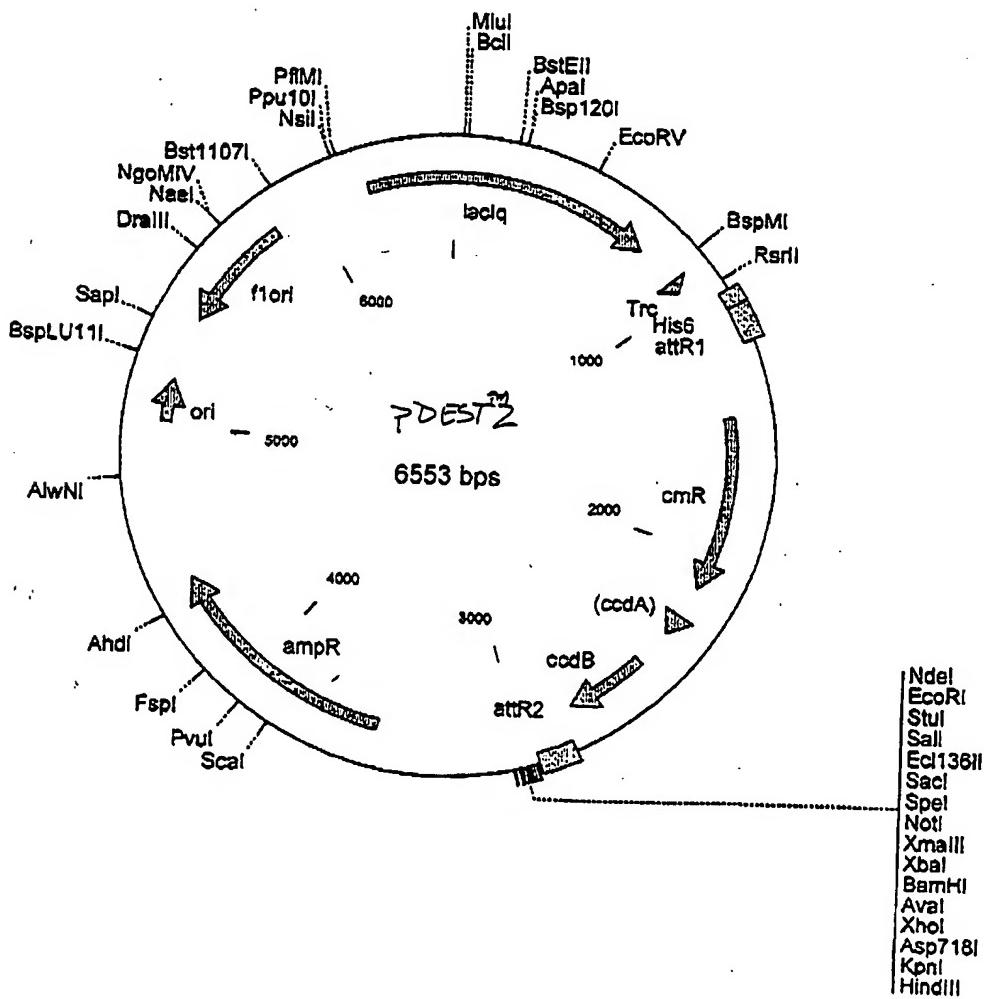
FIGURE 21D

36/240

Figure 22A: PDCST 2

His6 fusions in E. coli

970 aat att ctg aaa tga gct gtt gac att taa tca tcc ggt ccg cat aat ctg
 tta taa gac ttt act cga caa ctg tta att agt agg cca ggc ata tta gac
 1021 tgg aat tgt gag cgg ata aca att tca cac agg aaa cag acc atg teg ttc
 acc tta aca ctc gcc tat tgt taa agt gtg tcc ttt gtc tgg tac agc atg
 1072 Tyr His His His His His Gly Tle The S Tyt crt-R1
 tac cat ctc cat cat cat ctc ggt att aca agt tgg CAD xx aa gcy gaa
 atg gta gtg gta gtg gta gtg ccg tag tgt tca aac atg ttt crt cga crt



37/240

pDEST2 6553 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
912..962	Trc
1223..1009	attR1
1473..2132	CmR
2252..2336	inactivated ccdA
2474..2779	ccdB
2820..2944	attR2
3509..4414	ampR
5015..5175	ori
5415..5852	flori (f1 intergenic region)
6225..752	lacIq

1 GGCGGTGCAC AATCTTCTCG CGCAACCGTG CAGTGGGCTG ATCATTAACT ATCCGCTGGA
 61 TGACCAGGAT GCCATTGCTG TGGAAAGCTGC CTGCACTAAT GTTCCGGCGT TATTTCCTGA
 121 TGTCTCTGAC CAGACACCCA TCAACAGTAT TATTTCCTCC CATGAAGACG GTACGCGACT
 181 GGGCGTGGAG CATCTGGTGC CATTGGGTCA CCAGCAAATC GCGCTGTTAG CGGGCCCAT
 241 AAGTTCTGTC TCGGCGCGTC TGCCTCTGGC TGGCTGGCAT AAATATCTCA CTCGCAATCA
 301 AATTCAAGCCG ATAGCGGAAC GGGAAAGCGA CTGGAGTGCCT ATGTCGGTT TTCAACAAAC
 361 CATGCAAATG CTGAATGAGG GCATCGTTC CACTGGATG CTGGTTGCCA ACGATCAGAT
 421 GGCCTGGGGC GCAATGCGCG CCATTACCGA GTCCGGGCTG CGCCTGGGTG CGGATATCTC
 481 GGTAGTGGGA TACGACGATA CCGAAGACAG CTCATGTTAT ATCCCCTCCG CAACCAACAT
 541 CAAACAGGAT TTTCGCTGC TGGGGCAAAC CAGCGTGGAC CGCTTGTGCA AACTCTCTCA
 601 GGGCCAGGCG GTGAAGGGCA ATCAGCTGTT GCCCCCTCA CTGGTGA AAAAACCAC
 661 CCTGGCACCC AATACGAAA CCGCCTCTCC CGCGCGTTG CGCGATTCA TAATGCAGCT
 721 GGCACGACAG GTTTCGGAC TGGAAAGCGG GCAGTGAGCG CAACGCAATT AATGTGAGTT
 781 AGCGCGAATT GATCTGGTTT GACAGCTTAT CATCGACTGC ACGGTGCACC AATGCTTCTG
 841 GCGTCAGGCA GCCATCGGAA GCTGTGGTAT GGCTGTGCA GTCGTAATC ACTGCATAAT
 901 TCGTGTGCT CAAGGCGCAC TCCCGTTCTG GATAATGTTT TTTGCGCCGA CATCATAACG
 961 GTTCTGGAA ATATTCTGAA ATGAGCTGTT GACAATTAAAT CATCCGGTCC GTATAATCTG
 1021 TGGAAATTGTG AGCGGATAAC AATTTCACAC AGGAAACAGA CCATGTCGTA CTACCATCAC
 1081 CATCACCATC ACGGCATCAC AAGTTTGAC AAAAAAGCTG AACGAGAAAC GTAAAATGAT
 1141 ATAATATATCA ATATATTTAA TTAGATTTG CATAAAAAAC AGACTACATA ATACTGTAAA
 1201 ACACAACATA TCCAGTCACT ATGGCGCCCG CTAAGTTGGC AGCATCACCC GACGCACCTT
 1261 GCGCCGAATA AATACCTGTG ACAGGAAGATC ACTTCGAGA ATAAATAAAT CCTGGTGTCC
 1321 CTGTTGATAC CGGGAGGCC TGGGCAACT TTGGCGAAA ATGAGACGTT GATCGGCACG
 1381 TAAGAGGTTC CAACTTTCAC CATAATGAAA TAAGATCAGT ACCGGGCGTA TTTTTGAGT
 1441 TATCGAGATT TTCAGGAGCT AAGGAAGCTA AAATGGAGAA AAAATCACT GGATATACCA
 1501 CCGTTGATAT ATCCCAATGG CATCGTAAAG AACATTTGA GGCATTTCA TCAGTTGCTC
 1561 AATGTACCTA TAACCAGACG GTTCAGCTGG ATATTACGGC CTTTTAAAG ACCGTAAGA
 1621 AAAATAAGCA CAAGTTTAT CGGGCTTAA TTCACATTCT TGCCCGCTG ATGAATGCTC
 1681 ATCCGAATT CGGTATGGCA ATGAAAGACG GTGAGCTGGT GATATGGGAT AGTGTTCACC
 1741 CTTGTTACAC CGTTTCTCAT GAGCAAATG AAACGTTTTC ATCGCTCTGG AGTGAATACC
 1801 ACGACGATTT CGGGCAGTTT CTACACATAT ATTGCAAGA TGTGGCGTGT TACGGTGA
 1861 ACCTGGCTA TTCCCTAAAG GGGTTTATTG AGAATATGTT TTTCGTCCTCA GCGAACCTC
 1921 GGGTAGTTT CACCAAGTTT GATTTAAACG TGGCCAATAT GGACAACCTTC TTGCCCCCG
 1981 TTTCACCAT GGGCAAATAT TATACGCAAG GCGACAAGGT GCTGATGCCG CTGGCGATTC
 2041 AGGTTCATCA TGCGCTCTGT GATGGCTTCC ATGTCGGCAG AATGCTTAAAT GAATTACAAC
 2101 AGTACTGGCA TGAGTGGCA CGCGGGGGCTTAAACCGCTGG ATCCGGCTTA CTAAAAGCCA
 2161 GATAACAGTA TGGTATTG CGCGCTGATT TTTCGGTAT AAGAATATAT ACTGATATGT
 2221 ATACCCGAAG TATGTCAAAAG AGAGGTGTGC TATGAAGCAG CGTATTACAG TGACAGTTGA
 2281 CAGCGACAGC TATCAGTTGC TCAAGGCATA TATGATGTCA ATATCTCCGG TCTGGTAAGC
 2341 ACAACCATGC AGAATGAAGC CGTCGCTCTG CGTGCAGAC GCTGGAAAGC GGAAAATCAG
 2401 GAAGGGATGG CTGAGGTCGC CGGGTTTATT GAAATGAACG GCTCTTTGC TGACGAGAAC
 2461 AGGGACTGGT GAAATGCAGT TTAAGGTTA CACCTATAAA AGAGAGAGCC GTTATCGTCT
 2521 GTTGTGGAT GTACAGAGTG ATATTATTGA CACGCCGGG CGACGGATGG TGATCCCCCT-

FIGURE 22B

38/240

2581 GGCCAGTGCA CGTCTGCTGT CAGATAAAAGT CTCCC GTGAA CTTTACCCGG TGGTGCATAT
 2641 CGGGGATGAA AGCTGGCGCA TGATGACCAC CGATATGCC AGTGTGCCGG TCTCCGTTAT
 2701 CGGGGAAGAA GTGGCTGATC TCAGGCCACCG CGAAAATGAC ATCAAAAACG CCATTAACCT
 2761 GATGTTCTGG GGAATATAAA TGTCAAGGCTC CCTTATACAC AGCCAGTCTG CAGGTGACC
 2821 ATAGTGACTG GATATGTTGT GTTTTACAGT ATTATGTTAGT CTGTTTTTA TGCAAAATCT
 2881 AATTAAATAT ATTGATATTAT ATATCATTTC ACGTTTCTCG TTCAGCTTTC TTGTAACAAAG
 2941 TGGTGATGCC CATATGGAA TTCAAAGGCC TACGTCGACG AGCTCACTAG TCGCGGCCGC
 3001 TTCTAGAGGA TCCCTCGAGG CATGCGGTAC CAAGCTTGGC TGTTTGGCG GATGAGAGAA
 3061 GATTTCAGC CTTGATACAGA TTAAATCAGA ACGCCAGAAGC GGTCTGATAA AACAGAATT
 3121 GCCTGGCGC AGTAGCGCG TGTTCCCACC TGACCCCCATG CCGAACCTCAG AAGTGAACAG
 3181 CCGTAGCGCC GATGGTAGTG TGGGTTCTCC CCATGCGAGA GTAGGGAACT GCCAGGCATC
 3241 AAATAAAAAGC AAAGGCTCAG TCGAAAGACT GGGCCTTTCG TTTTATCTGT TGTTTGTGG
 3301 TGAAACGCTCT CTCAGTAGG ACAAAATCCG CGGGAGCGGA TTGAAACGTT CGGAAGCAAC
 3361 GGCCCGGAGG GTGGCGGGCA GGACGCCGC CATAAAACTGC CAGGCATCAA ATTAAGCAGA
 3421 AGGCCATCCT GACGGATGGC TTTCCTGCGT TTCTACAAAC TCTTTTTGTT TATTTTCTA
 3481 AATACATTCA AATATGTATC CGCTCATGAG ACAATAACCC TGATAAATGC TTCAATAATA
 3541 TTGAAAAAGG AAGAGTATGA GTATTCAACA TTTCCTGTC GCCCTTATTTC CCTTTTTGTC
 3601 GGCATTTGC CTTCTGTTT TTGCTCACCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA
 3661 AGATCAGTTG GGTGCACGAG TGGGTTACAT CGAACCTGGAT CTCAACAGCG GTAAGATCCT
 3721 TGAGAGTTT CGCCCCGAAG AACGTTTCC AATGATGAGC ACTTTAAAG TTCTGCTATG
 3781 TGGCGCGGTAA TTATCCCGTG TTGACGCCGG GCAAGAGCAA CTCGGTCGCC GCATACACTA
 3841 TTCTCAGAAT GACTTGGTT AGTACTCACC AGTCACAGAA AAGCATCTT CCGATGGCAT
 3901 GACAGTAAGA GAATTATGCA GTGCTGCCAT ACCATGAGT GATAACACTG CGGCCAACTT
 3961 ACTTCTGACA ACGATCGGAG GACCGAAGGA GCTAACCGCT TTTTGCAAA ACATGGGGGA
 4021 TCATGTAACT CGCCTTGATC GTTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGACGA
 4081 GCGTGACACC ACGATGCCA CAGCAATGGC AACACGTTG CGCAAACATAT TAATGGCGA
 4141 ACTACTTACT CTAGCTTCCC GGCAACAATT AATAGACTGG ATGGAGGCAG ATAAAGTTGC
 4201 AGGACCACCTT CTGCGCTCGG CCCTTCCGGC TGGCTGGTTT ATTGCTGATA AATCTGGAGC
 4261 CGGTGAGCGT GGGTCTCGCG GTATCATTGC AGCACTGGGG CCAGATGGTA AGCCCTCCG
 4321 TATCGTAGTT ATCTACACGA CGGGGAGTCA GGCAACTATG GATGAAAGAA ATAGACAGAT
 4381 CGCTGAGATA GGTGCCTCAC TGATTTAACCA TTGCTTAACAG TCAGACCAAG TTTACTCATA
 4441 TATACTTTAG ATTGATTTAA AACTTCATTT TTAATTTAAA AGGATCTAGG TGAAGATCCT
 4501 TTTTGATAAT CTCATGACCA AAATCCCTTA ACGTGAGTTT TCGTTCCACT GAGCGTCAGA
 4561 CCCCGTAGAA AAGATCAAAG GATCTTCTTG AGATCCTTTT TTCTGCGCG TAATCTGCTG
 4621 CTTGCAAACCA AAAAACCAC CGCTACCCAGC GGTGGTTTGT TTGCGGGATC AAGAGCTACC
 4681 AACTTTTTT CCGAAGGTAA CTGGCTTCAG CAGAGCGAG ATACCAAATA CTGTCCTTCT
 4741 AGTGTAGCCG TAGTTAGGCC ACCACTTCAA GAACTCTGTA GCACCGCCTA CATACTCGC
 4801 TCTGCTAATC CTGTTACCAAG TGGCTCTGC CAGTGGCGAT AGTCTGTC TTACCGGGTT
 4861 GGACTCAAGA CGATAGTTAC CGGATAAGC GCAGCGGTG GGCTGAACGG GGGGTTCTG
 4921 CACACAGCCC AGCTTGGAGC GAACGACCTA CACCGAACTG AGATACCTAC AGCGTGAGCT
 4981 ATGAGAAAGC GCCACGCTTC CGCAAGGGAG AAAGGGGGAC AGGTATCCGG TAAGCGGAG
 5041 GGTGGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGGAA AACGCCCTGGT ATCTTATAG
 5101 TCCCTGCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTTGATGCT CGTCAGGGGG
 5161 GCGGAGCCTA TGGAAAAAACCG CCAGCAACCG GGCCTTTTA CGGTTCTGG CCTTTGCTG
 5221 GCCTTTGCT CACATGTTCT TTCCCTGCGT ATCCCTGAT TCTGTTGATA ACCGTATTAC
 5281 CGCCTTGAG TGAGCTGATA CCGCTCGCCG CAGCCGAACG ACCGAGCGCA GCGAGTCAGT
 5341 GAGCGAGGAA GCGGAAGAGC GCCTGATGCG GTATTTCTC CTTACGCATC TGTGCGGTAT
 5401 TTCACACCGC ATAATTTTG TAAAATTGCG GTTAAATTTT TGTTAAATCA GCTCATTCTT
 5461 TAACCAATAG GCGAAATCG GCAAATCCC TTATAAATCA AAAGAATAGA CCGAGATAGG
 5521 GTTGAGTGTT GTTCCAGTTT GGAACAAGAG TCCACTATTA AAGAACGTGG ACTCCAACGT
 5581 CAAAGGGCGA AAAACCGTCT ATCAGGGCGA TGGCCCACTA CGTGAACCAT CACCCATAATC
 5641 AAGTTTTTG GGGTCGAGGT GCGTAAAGC ACTAAATCGG AACCCCTAAAG GGAGCCCCCG
 5701 ATTTAGAGCT TGACGGGGAA AGCCGGCAGA CGTGGCGAGA AGGAAGGGAGA AGAAAGCGAA
 5761 AGGAGCGGGC GCTAGGGCGC TGGCAAGTGT AGCGGTCACG CTGCGCGTAA CCACCAACCC
 5821 CGCGCGCTT AATGCGCCGC TACAGGGCGC GTCCCATTCG CCATTCAGGC TGCTATGGTG
 5881 CACTCTCAGT ACAATCTGCT CTGATGCCGC ATAGTTAACG CAGTATACAC TCCGCTATCG
 5941 CTACGTGACT GGGTCATGCC TGCGCCCCGA CACCCGCCAA CACCCGCTGA CGCGCCCTGA
 6001 CGGGCTTGTC TGCTCCCGGC ATCCGTTAC AGACAAGCTG TGACCGTCTC CGGGAGCTGC-

FIGURE 22C

39/240

6061 ATGTGTCAGA GGTTTTCACC GTCATCACCG AAACGCCGA GGCAGCAGAT CAATTGCGC
6121 GCGAAGGCCGA AGCGGCATGC ATTTACGTTG ACACCATCGA ATGGTGCAA ACCTTCGCG
6181 GTATGGCATG ATAGCGCCCG GAAGAGAGTC AATTCAAGGGT GGTGAATGTG AAACCAGTAA
6241 CGTTATACGA TGTCGCAGAG TATGCCGGTG TCTCTTATCA GACCGTTCC CGCGTGGTGA
6301 ACCAGGCCAG CCACGTTCT GCGAAAACGC GGGAAAAAGT GGAAGCGCCG ATGGCGGAGC
6361 TGAATTACAT TCCCAACCGC GTGGCACAAAC AACTGGCGGG CAAACAGTCG TTGCTGATTG
6421 GCGTTGCCAC CTCCAGTCTG GCCCTGCACG CGCCGTCGCA AATTGTCGCG GCGATTAAT
6481 CTCGCGCCGA TCAACTGGGT GCCAGCGTGG TGGTGTGAT GGTAGAACGA AGCGGGCGTCG
6541 AAGCCTGTAA AGC

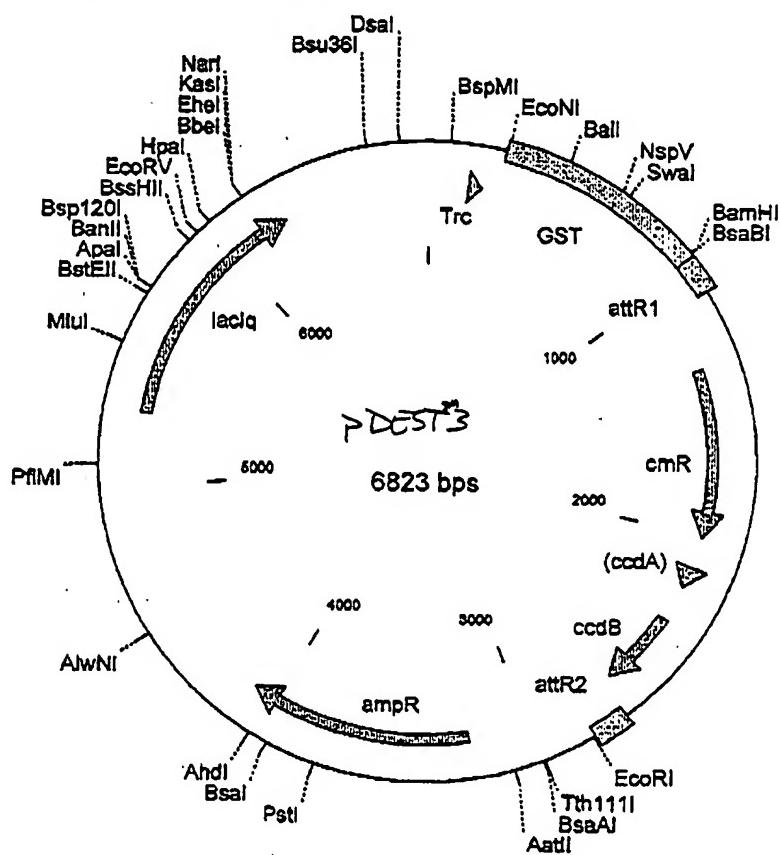
FIGURE 22D

40/240

Figure 23A: PDEST3

GST fusions in *E. coli*

154 cggttc tggcaa atatcc tga aat gag ctg -15 Trc promoter
 gcc aag acc gtt tat aag act tta ctc gac aac tgt taa tta gta gec gag
 .205 gta taa tgt gtg gaa ttg tga gcg gat aac aat ttc aca cag gaa aca gta
 cat att aca cac ctt aac act cgc cta ttg tta aag tgt gtc ctt tgt cat
 256 ttc atg tcc cct ata cta ggt tat tgg aaa att aag ggc ctt gtg caa ccc
 aag tac agg gga tat gat cca ata acc ttt taa ttc ccg gaa cac gtt ggg



41/240

pDEST3 6823 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
150..200	Trc
1087..963	attR1
1337..1996	CmR
2116..2200	inactivated ccda
2338..2643	ccdB
2684..2808	attR2
3231..4091	ampR
5295..6254	lacIq

1 ACGTTATCGA CTGCACGGTG CACCAATGCT TCTGGCGTCA GGCAGCCATC GGAAGCTGTG
 61 GTATGGCTGT GCAGGGTCGTA AATCACTGCA TAATTCGTGT CGCTCAAGGC GCACTCCCGT
 121 TCTGGATAAT GTTTTTGCG CGGACATCAT AACGGTTCTG GCAAATATTG TGAAATGAGC
 181 TGTTGACAAT TAATCATCGG CTCGTATAAT GTGTTGAAATT GTGAGCGGAT AACAAATTCA
 241 CACAGGAAAC AGTATTCTAG TCCCCCTATAC TAGGTTATTG GAAAAATTAAAG GGCCTTGTGC
 301 AACCCACTCG ACTTCTTTTG GAATATCTG AAGAAAATA TGAAGAGCAT TTGATGAGC
 361 GCGATGAAGG TGATAAATGG CGAAACAAAA AGTTTGAATT GGTTTGGAG TTTCCCACATC
 421 TTCCTTATTA TATTGATGGT GATGTTAAAT TAACACAGTC TATGCCATC ATACGTTATA
 481 TAGCTGACAA GCACAAACATG TTGGGTTGGTT GTCCAAAAGA GCGTGCAGAG ATTTCAATGC
 541 TTGAAGGAGC GTTTTGGAT ATTAGATACG GTGTTTCGAG AATTGCAATAT AGTAAAGACT
 601 TTGAAACTCT CAAAGTTGAT TTTCTTAGCA AGTACCTGA AATGCTGAAA ATGTTGAAAG
 661 ATCGTTTATG TCATAAAAACA TATTTAAATG GTGATCATGT AACCCATCCT GACTTCATGT
 721 TGTATGACGC TCTTGATGTT GTTTTATACA TGGACCCAAT GTGCCCTGGAT GCGTTCCCAA
 781 AATTAGTTG TTTAAAAAAA CGTATTGAAAG CTATCCCACA AATTGATAAAG TACTTGAAAT
 841 CCAGCAAGTA TATAGCATGG CCTTTGCGAG GCTGGCAAGC CACGTTGGT GGTGGCGACC
 901 ATCCTCCAAA ATCGGATCTG GTTCCGGTGT GATCTCGTCG TGCATCTGTT GGATCCCCAT
 961 CAACAAGTTT GTACAAAAAA GCTGAACGAG AAACGTAAAA TGATATAAAT ATCAATATAT
 1021 TAAATTAGAT TTTGCATAAA AAACAGACTA CATAAACTG TAAAACACAA CATATCCAGT
 1081 CACTATGGCG GCGCCTAAGT TGGCAGCATC ACCCGACGCA CTTTGCGCCG AATAAAATACC
 1141 TGTGACGGAA GATCACCTCG CAGAATAAAT AAATCCTGGT GTCCCTGTTG ATACCGGGAA
 1201 GCCCTGGGCC AACTTTGGC GAAAATGAGA CGTTGATCGG CACGTAAGAG GTTCCAACCTT
 1261 TCACCCATAAT GAAATAAGAT CACTACCGG CGTATTTTT GAGTTATCGA GATTTTCAGG
 1321 AGCTAAGGAA GCTAAAATGG AGAAAAAAAT CACTGGATAT ACCACCGTTG ATATATCCCA
 1381 ATGGCATTGCG AAAGAACATT TTGAGGCAATT TCAGTCAGTT GCTCAATGTA CCTATAACCA
 1441 GACCGTTCAAG CTGGATATTA CGGCCCTTTT AAAGACCGTA AAGAAAAATA AGCACAAGTT
 1501 TTATCCGGCC TTTATTCAAC TTCTTGGCCG CCTGATGAAT GCTCATCCGG AATTCCGTAT
 1561 GGCAATGAAA GACGGTGAGC TGGTGTATG GGATAGTGTGTT CACCCCTGTT ACACCGTTTT
 1621 CCATGAGCAA ACTGAAACGTT TTTCATCGCT CTGGAGTGAA TACCAACGACG ATTTCCGGCA
 1681 GTTTCTACAC ATATATTGCG AAGATGTGGC GTGTTACGGT GAAAACCTGG CCTATTCCC
 1741 TAAAGGGTTT ATTGAGAATA TCTGATGTTCTGCAAT CCTGGGTTGA GTTTCACCAAG
 1801 TTTTGATTAA AACGTGGCCA ATATGGACAA CTTCTTCGCC CCCGTTTCA CCATGGCAA
 1861 ATATTATACG CAAGGCAGACA AGGTGCTGAT GCGGCTGGCG ATTCAGGTTC ATCATGCCGT
 1921 CTGTGATGGC TTCCATGTCG GCAGAAATGCT TAATGAATTAA CAACAGTACT GCGATGAGTG
 1981 GCAGGGCGGG CGCTAAAGAT CTGGATCCGG CTTACTAAAA CCCGAGATAAAC AGTATGCCA
 2041 TTTGCGCGCT GATTTTTGCG GTATAAGAAT ATATACTGAT ATGTATACCC GAAGTATGTC
 2101 AAAAAGAGGT GTGCTATGAA GCAGCGTATT ACAGTGCACAG TTGACAGCGA CAGCTATCAG
 2161 TTGCTCAAGG CATATATGAT GTCAATATCT CCGGTCTGGT AACGACAACC ATGCAGAATG
 2221 AAGCCCGTCG TCTGCGTGC GAACGCTGGA AAGCGAAAAA TCAGGAAGGG ATGGCTGAGG
 2281 TCGCCCGGTT TATTGAAATG AACGGCTCTT TTGCTGACGA GAACAGGGAC TGGTGAATG
 2341 CAGTTTAAGG TTTACACCTA TAAAAGAGAG AGCCGTTATC GTCTGTTTGT GGATGTACAG
 2401 AGTGTATTTA TTGACACGCC CGGGCGACGG ATGGTGTATCC CCCTGGCCAG TGCACGTCTG
 2461 CTGTCAGATA AAGTCTCCCG TGAACATTAC CCGGTGGTGC ATATCGGGGA TGAAAGCTGG
 2521 CGCATGATGA CCACCGATAT GGCCAGTGTG CCGGTCTCCG TTATCGGGGA AGAAGTGGCT
 2581 GATCTCAGCC ACCGCGAAAAA TGACATCAA AACGCCATTA ACCTGATGTT CTGGGGAAATA
 2641 TAAATGTCAG GCTCCCTTAT ACACAGCCAG TCTGCGAGTC GACCATAGTG ACTGGATATG-

FIGURE 23B

42/240

2701 TTGTGTTTA CAGTATTATG TAGTCTGTT TTTATGAAA ATCTAATTAA ATATATTGAT
 2761 ATTTATATCA TTTTACGTT CTCGTTCAAC TTTCTTGAC AAAGTGGTTG ATGGGAATTG
 2821 ATCGTACTG ACTGACGATC TGCCCTCGGC GTTTCGGTGA TGACGGTGAA AACCTCTGAC
 2881 ACATGCAGCT CCCGGAGAGC GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAG
 2941 CCCGTCAGGG CGCGTCAGCG GGTGTTGGCG GGTGTCGGGG CGCAGCCATG ACCCAGTCAC
 3001 GTAGCGATAG CGGAGTGTAT AATTCTGAA GACGAAAGGG CCTCGTATA CGCCTATTG
 3061 TATAGGTTAA TGTCATGATA ATAATGTTT CTTAGACGTC AGGTGGCACT TTTCGGGAA
 3121 ATGTGCGCGG AACCCCTATT TGTTTATTT TCTAAATACA TTCAAATATG TATCCGCTCA
 3181 TGAGACAAATA ACCCTGATAA ATGCTTCAT AATATTGAAA AAGGAAGAGT ATGAGTATTG
 3241 AACATTTCCG TGTCGCCCTT ATTCCCCTTT TTGCGGCATT TTGCGCTTCCT GTTTTGCTC
 3301 ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTGGGGTGC CGAGTGGTT
 3361 ACATCGAACT GGATCTAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT
 3421 TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTGTGACG
 3481 CGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT
 3541 CACCACTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGC
 3601 CCATAACCAC AGGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA
 3661 AGGAGCTAAC CGCTTTTTG CACAACATGG GGGATCATGT AACTCGCCCTT GATCGTTGGG
 3721 AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGCAGCAA
 3781 TGGCAACAAAC GTTGCGBAA CTATTAACG GCAGAACTACT TACTCTAGCT TCCCGGCAAC
 3841 AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAAGGACC ACTTCTGCGC TCGGCCCTC
 3901 CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA
 3961 TTGCAAGCACT GGGGCCAGAT GGTAAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGG
 4021 GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA
 4081 AGCATTGGTA ACTGTCAGAC CAAGTTACT CATATATACT TTAGATTGAT TTAAACCTTC
 4141 ATTTTAATT TAAAAGGATC TAGGTGAAGA TCCTTTTGA TAATCTCATG ACCAAATTC
 4201 CTTAACGTGA GTTTCTGTC CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT
 4261 CTTGAGATCC TTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCGCTAC
 4321 CAGCGGTGGT TTGTTGCGG GATCAAGAGC TACCAACTCT TTTTCGAAAG GTAACGGCT
 4381 TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTG GCCGTAGTTA GGCCACCACT
 4441 TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT AATCTGTTA CCAGTGGCTG
 4501 CTGCCAGTGG CGATAAGTCG TGTCTTACCG GGTTGGACTC AAGACGATAG TTACCGGATA
 4561 AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGCACACA GCCCAGCTTG GAGCGAACGA
 4621 CCTACACCGA ACTGAGATAAC CTACAGCGTG AGCTATGAGA AAGGCCACCG CTTCCCGAAG
 4681 GGAGAAAGGC GGACAGGTAT CGGTAAAGCG GCAGGGTCGG AACAGGAGAG CGCACGAGGG
 4741 AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC
 4801 TTGAGCGTCG ATTTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA
 4861 ACGGCGGCTT TTACGGTTC CTGGCTTTT GCTGGCTTT TGCTCACATG TTCTTCTG
 4921 CGTTATCCCC TGATTCTGAG GATAACCGTA TTACCGCTT TGAGTGGAGCT GATACCGCTC
 4981 GCCCGAGCCG AACGACCGAG CGCAGCGAGT CAGTGGCGA GGAAGCGGAA GAGCGCTCTGA
 5041 TGCGGTATTT TCTCTTACG CATCTGTGCG GTATTTACA CGCGATAAAAT TCCGACACCA
 5101 TCGAATGGTG CAAAACCTTT CGCGGTATGG CATGATAGCG CCCGGAAGAG AGTCAATTCA
 5161 GGGTGGTGAA TGTGAAACCA GTAACTGTT ACGATGTCGC AGAGTATGCC GGTGTCCTT
 5221 ATCAGACCGT TTCCCACG TGGAACCGG CCAGCCACGT TTCTGCGAAA ACGCCGGAAA
 5281 AAGTGGAAAGC GGCGATGGCG GAGCTGAATT ACATCCCCA CGCGGTGGCA CAACAACTGG
 5341 CGGGCAACAA GTCGTTGCTG ATTGGCGTT CCACCTCCAG TCTGGCCCTG CACGCCCGT
 5401 CGCAAATTGT CGCGCGATT AAAATCTCGC CGCATCACT GGGTGCCACG GTGGTGGTGT
 5461 CGATGGTAGA ACGAAGCGGC GTCGAAGCCT GTAAAGCGGC GGTGACAAT CTTCTCGCGC
 5521 AACCGTCAG TGGGCTGATC ATTAACATAC CGCTGGATGA CCAGGATGCC ATTGCTGTGG
 5581 AAGCTGCTG CACTAATGTT CCGCGTTT TTCTGTGTT CTCTGACCAAG ACACCCATCA
 5641 ACAGTATTAT TTCTCCCAT GAAGACCGTA CGCGACTGGG CGTGGAGCAT CTGGTGCAT
 5701 TGGGTACCCA GCAAATCGCG CTGTTACCGG GCCCCATTAAG TTCTGTCTCG GCGCGCTCTGC
 5761 GTCTGGCTGG CTGGCATAAA TATCTCACTC GCAATCAAAT TCAGCCGATA GCGGAACGGG
 5821 AAGGCGACTG GAGTGGCCATG TCCGGTTTCA AACAAACCAT GAAATGCTG AATGAGGGCA
 5881 TCGTTCCCAC TGCGATGCTG GTTGCACAGC ATCAGATGGC GCTGGCGCA ATGCGCGCCA
 5941 TTACCGAGTC CGGGCTGCGC GTTGGTGCAGG ATATCTCGGT AGTGGGATAC GACGATAACCG
 6001 AAGACAGCTC ATGTTATATC CCGCCGTTAA CCACCATCAA ACAGGATTTT CGCCTGCTGG
 6061 GGAAACCAAG CGTGGACCGC TTGCTGAAAC TCTCTCAGGG CCAGGGCGGTG AAGGGCAATC
 6121 AGCTGTTGCC CGTCTCACTG GTGAAAGAA AAACCCACCTT GCGCCCAAT ACGCAAACCG-

FIGURE 23C

43/240

6181 CCTCTCCCCG CGCGTTGGCC GATTCAATTAA TGCAGCTGGC ACGACAGGTT TCCCAGCTGG
6241 AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC TCACTCATTA GGCACCCAG
6301 GCTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATT
6361 CACACAGGAA ACAGCTATGA CCATGATTAC GGATTCACTG GCCGTCGTTT TACAACGTCG
6421 TGACTGGAA AACCCCTGGCG TTACCCAATC TAATCGCCTT GCAGCACATC CCCCTTCGC
6481 CAGCTGGCGT AATAGCGAAG AGGCCCGCAC CGATCGCCCT TCCCAACAGT TGCGCAGCCT
6541 GAATGGCGAA TGGCGCTTTG CCTGGTTTCC GGCACCAGAA CGGGTGCCGG AAAGCTGGCT
6601 GGAGTGCAGAT CTTCTGAGG CCGATACTGT CGTCGTCCCC TCAAACTGGC AGATGCACGG
6661 TTACCGATGCG CCCATCTACA CCAACGTAAC CTATCCCATT ACGGTCAATC CGCCGTTTGT
6721 TCCCCACGGAG AATCCGACGG GTTGTACTC GCTCACATTT AATGTTGATG AAAGCTGGCT
6781 ACAGGAAGGC CAGACGCGAA TTATTTTGA TGGCGTTGGA ATT

FIGURE 23D

44/240

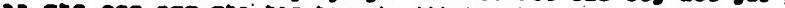
Figure 24A: pDEST4

His6-thioredoxin fusions in E. coli

-35 Tre promoter -10
919 gca aat att ctg aaa tga get ggt gac att taa tca tcc ggt ccg cat aat
cgt tta taa gac ttt act cga cca ctg tca att agt agg cca ggc tta tca

978 ctg tgg ~~aat~~^{taat} tgt gag cgg ata.aca att tea cac agg aaa cag acc ~~tgt~~^{tac} gtc
gac acc tta aca ctc gcc tat tgt taa agt gtg tcc ttt gtc tgg tac cca

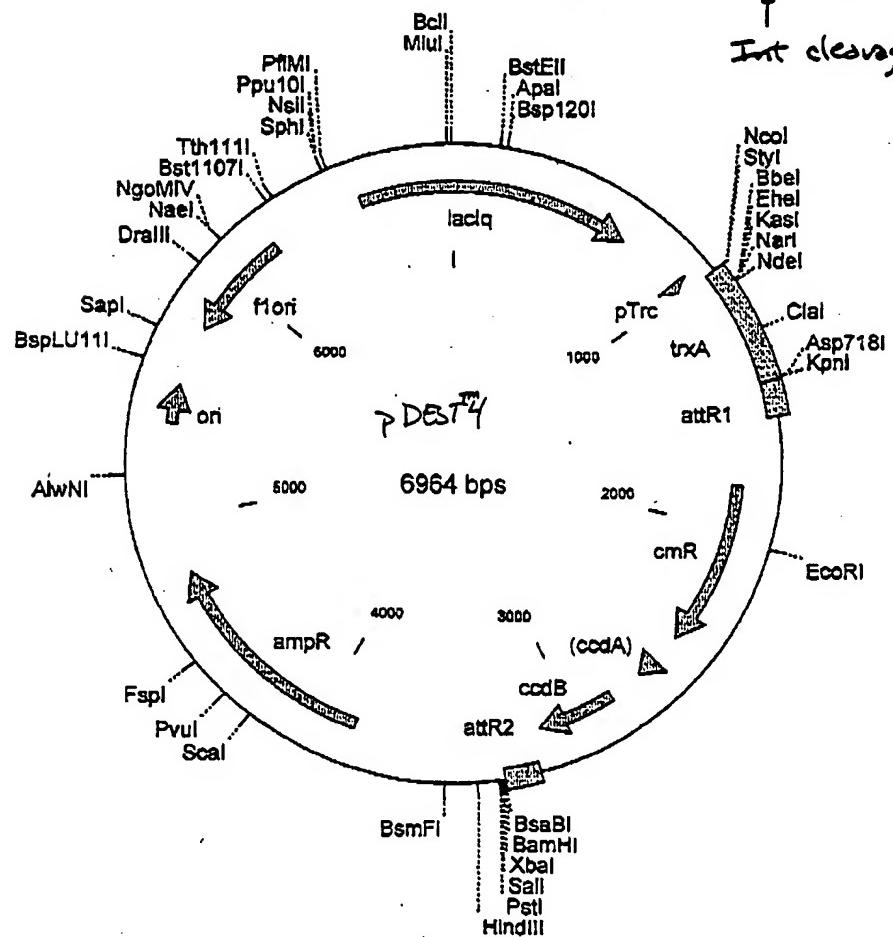
Hs 6

1072 TEV protease → Thioredoxin - - (≈ 150 amino acids)


```

graph LR
    A[TEV protease] --> B[Thioredoxin]
    A --> C[Thioredoxin]
    
```

1429 A₁ A₂ A₃ A₄ Lys Val Pro Ile Tyr Sec Leu Thr Lys Lys
gat gac gat gat gtc aag gta ccc atc tca aat tgc tcc tcc tcc tcc tcc tcc
cta ctg cta ctg ttc cat ggg tag tct tca aac aac ttt ttt tga pcc act



45/240

pDEST4 6964 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
964..1003	Trc
1577..1453	attR1
1827..2486	CmR
2606..2690	inactivated ccdA
2828..3133	ccdB
3174..3298	attR2
3872..4777	ampR
5378..5538	ori
5778..6215	flori (f1 intergenic region)
6587..704	lacIq

1 CTATCCGCTG GATGACCAGG ATGCCATTGC TGTGGAAGCT GCCTGCACTA ATGTTCCGGC
 61 GTTATTTCTT GATGTCCTCG ACCAGACACC CATCAACAGT ATTATTTCT CCCATGAAGA
 121 CGGTACCGA CTGGCGTGG AGCATCTGGT CGCATTGGGT CACCAAGCAA TCGCGCTGTT
 181 AGCGGGCCA TTAAGTTCTG TCTCGCGCG TCTCGCTCTG GCTGGCTGGC ATAAATATCT
 241 CACTCGCAAT CAAATTCAAG CGATAGCGGA ACGGGAAGGC GACTGGAGTG CCATGTCGG
 301 TTTCAACAA ACCATGCAA TGCTGAATGA GGGCATCGTT CCCACTGCGA TGCTGGTTGC
 361 CAACGATCAG ATGGCGCTGG GCGCAATGCG CGCCATTACG GAGTCCGGGC TGCGCGTTGG
 421 TGCGGATATC TCAGGTAGTGG GATAACGACG TACCGAAGAC AGCTCATGTT ATATCCGCC
 481 GTCAACCACC ATCAAACAGG ATTTTCGCCT GCTGGGGCAA ACCAGCGTGG ACCGCTTGCT
 541 GCAACTCTCT CAGGGCCAGG CGGTGAAGGG CAATCAGCTG TTGCCCCGTCT CACTGGTGAA
 601 AAGAAAAACC ACCCTGGCAC CCAATACGCA AACCGCCTCT CCCCAGCGGT TGGCCGATT
 661 ATTAATGCAG CTGGCACGAC AGGTTCCCG ACTGGAAAGC GGGCAGTGAG CGCAACGCAA
 721 TTAATGTGAG TTAGCGCGAA TTGATCTGGT TTGACAGCTT ATCATCGACT GCACGGTGCA
 781 CCAATGCTTC TGGCGTCAGG CAGCCATCGG AAGCTGTGGT ATGGCTGTGC AGGTCGTAAA
 841 TCACTGCATA ATTCTGTGTCG CTCAAGGCGC ACTCCCGTTC TGATAATGT TTTTGCGCC
 901 GACATCATAA CGGTTCTGGC AAATATTCTG AAATGAGCTG TTGACAATTAA ATCATCCGGT
 961 CCGTATAATC TGTGGAATTG TGAGCGGATA ACAATTTAAC ACAGGAAACA GACCATGGGT
 1021 CATCATCATC ATCATCACGA TTACGATATC CCAACGACCG AAAACCTGTA TTTTCAGGGC
 1081 GCCCATATGA GCGATAAAAT TATTACCTG ACTGACGACA GTTTTGACAC GGATGTACTC
 1141 AAAGCGGACG GGGCGATCCT CGTCGATTC TGGCGAGAGT GTGCGGGTCC GTGCAAATG
 1201 ATCGCCCGA TTCTGGATGA AATCGCTGAC GAATATCAGG GCAAAACTGAC CGTTGCAAAA
 1261 CTGAACATCG ATCAAAACCC TGGCACTGCC CCGAAATATG GCATCCGTGG TATCCCAGCT
 1321 CTGCTGCTGT TCAAAAACGG TGAAGTGGCG GCAACCAAAG TGGGTGCAC GTCTAAAGGT
 1381 CAGITGAAAG AGTTCTCGA CGCTAACCTG GCGCGTTCTG GTTCTGGTGA TGACGATGAC
 1441 AAGGTACCCA TCACAAGTTT GTACAAAAAA GCTGAACGAG AAACGTAAAAA TGATATAAAT
 1501 ATCAATATAT TAAATTAGAT TTTGCTAAA AAACAGACTA CATAATACTG TAAAACACAA
 1561 CATATCCAGT CACTATGGCG GCCGCTAAGT TGGCAGCATC ACCCGACGCA CTTTGCAGCG
 1621 AATAAATACC TGTGACGGAA GATCACTTCG CAGATAAAAT AAATCCTGGT GTCCCTGTTG
 1681 ATACCGGAA GCCCTGGGC AACTTTGGC GAAAATGAGA CGTTGATCGG CACGTAAGAG
 1741 GTTCCAACCTT TCACCATAAT GAAATAAGAT CACTACGGG CGTATTTTTT GAGTTATCGA
 1801 GATTTTCAGG AGCTAAGGAA GCTAAAATGG AGAAAAAAAT CACTGGATAT ACCACCGTTG
 1861 ATATATCCC ATGGCATCGT AAAGAACATT TTGAGGCATT TCAGTCAGTT GCTCAATGTA
 1921 CCTATAACCA GACCGTTCA CGGGATAATTAA CGGCCCTTTT AAAGACCGTA AAGAAAAATA
 1981 AGCACAAGTT TTATCCGGCC TTTATTCCAA TTCTTGGCCCG CCTGATGAAT GCTCATCCGG
 2041 AATTCCGTAT GGCAATGAAA GACGGTGAGC TGGTGATATG GGATAGTGTT CACCCCTGTT
 2101 ACACCGTTT CCATGAGCAA ACTGAAACGT TTTCATCGCT CTGGAGTGGAA TACCACGACG
 2161 ATTTCCGGCA GTTTCTACAC ATATATTCCG AAGATGTGGC GTGTTACGGT GAAAACCTGG
 2221 CCTATTTCCC TAAAGGGTTT ATTGAGAATA TGTTTTCTGT CTCAGCCAAT CCCTGGGTGA
 2281 GTTCCACCAG TTTGATTTA AACGTGGCCA ATATGGACAA CTTCTTCGCC CCCGTTTCA
 2341 CCATGGCAA ATATTATACG CAAGGCGACA AGGTGCTGAT GCGCGTGGC ATTCAAGGTT
 2401 ATCATGCCGT CTGTGATGGC TTCCATGTGCG GCAGAATGCT TAATGAATTA CAACAGTACT
 2461 GCGATGAGTG GCAGGGCGGG GCGTAAACGC GTGGATCCGG CTTACTAAAA GCCAGATAAC
 2521 AGTATGCGTA TTTGCGCGCT GATTTTGCG GTATAAGAAT ATATACTGAT ATGTATAACCC-

FIGURE 24B

46/240

2581 GAAGTATGTC AAAAAGAGGT GTGCTATGAA GCAGCGTATT ACAGTGACAG TTGACAGCGA
 2641 CAGCTATCAG TTGCTCAAGG CATATATGAT GTCAATATCT CGGGCTGGT AAGCACAACC
 2701 ATGCAGAATG AAGCCCGTGC TCTGCGTGC GAACGCTGGA AAGCGGAAAA TCAGGAAGGG
 2761 ATGGCTGAGG TCGCCCGGTT TATTGAAATG AACGGCTCTT TTGCTGACGA GAACAGGGAC
 2821 TGGTCAAATG CAGTTTAAGG TTTACACCTA TAAAAGAGAG AGCCGTTATC GTCTGTTGT
 2881 GGATGTACAG AGTGTATTA TTGACACGCC CGGGCGACGG ATGGTGTAC CCCTGGCCAG
 2941 TGCACGTCTG CTGTCAGATA AAGTCTCCC TGAACTTAC CGGGTGGTGC ATATCAGGGGA
 3001 TGAAAGCTGG CGCATGATGA CCACCGATAT GGCCAGTGTG CGGGTCTCCG TTATCAGGGGA
 3061 AGAAGTGGCT GATCTCAGCC ACCCGAAAA TGACATCAA AAGGCCATTA ACCTGATGTT
 3121 CTGGGAATA TAAATGTCAG GCTCCCTTAT ACACAGCCAG TCTGCAAGTC GACCATAGTG
 3181 ACTGGATATG TTGTGTTTA CAGTATTATG TAGTCTGTT TTTATGCAA ATCTAATTAA
 3241 ATATATTGAT ATTATATATCA TTTTACGTT CTCGTTCAAGC TTTCTGTAC AAAGTGGTGA
 3301 TGGGATCCT CTAGAGTCGA CCTGCAGTAA TCGTACAGGG TAGTACAAAT AAAAAGGCA
 3361 CGTCAGATGA CGTGCCTTTT TTCTTGTGAG CAGTAAGCTT GGCTGTTTG GCGGATGAGA
 3421 GAAGATTTTC AGCCTGATAC AGATTAATC AGAACGCAGA AGCGGTCTGA TAAAACAGAA
 3481 TTGCGCTGGC GGCAGTAGCG CGGGTGGCC ACCTGACCCCC ATGCCGAACCT CAGAAGTGA
 3541 ACGCCGTAGC GGCAGATGGTA GTGTGGGTC TCCCCATGCC AGAGTAGGGAA ACTGCCAGGC
 3601 ATCAAATAAA ACGAAAGGCT CAGTCGAAAG ACTGGGCCCT TGTTTTATC TGTTGTTGT
 3661 CGGTGAACGC TCTCCTGAGT AGGACAAATC CGCCGGGAGC GGATTTGAAC GTTGCAGAAGC
 3721 AACGGCCCGG AGGGTGGCG GCAGGAGGCC CGCCATAAAC TGCCAGGCAT CAAATTAAAGC
 3781 AGAAGGCCAT CCTGACGGAT GGCTTTTG CGTTTCTACA AACTCTTTTG GTTTATTTT
 3841 CTAATAACAT TCAAATATGT ATCCGCTCAT GAGACAATAA CCCTGATAAA TGCTTCAATA
 3901 ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT GTGCCCTTAA TTCCCTTTT
 3961 TGCGGCATTT TGCCTTCCTG TTTTGTCTCA CCCAGAAACG CTGGTGAAG TAAAAGATGC
 4021 TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT
 4081 CCTTGAGAGT TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTA AAGTTCTGCT
 4141 ATGTGGCGCG GTATTATCCC GTGTTGACGC CGGGCAAGAG CAAACTCGGTC GCCGCATAACA
 4201 CTATTCTCAG ATGACTTGG TTGAGTACTC ACCAGTCACA GAAAAGCATC TTACGGATGG
 4261 CATGACAGTA AGAGAATTAT GCAGTGTGCA CATAACCATG AGTGATAACA CTGCGGCCAA
 4321 CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACCG GCTTTTTGC ACAACATGGG
 4381 GGATCATGTA ACTCGCCTTG ATCGTTGGG ACCGGAGCTG AATGAAGCCA TACCAAACGA
 4441 CGAGCGTGAC ACCACGATGC CTACAGCAAT GGCAACAACG TTGCGAAAC TATTAACCTGG
 4501 CGAACTACTT ACTCTAGCTT CCCGCAACA ATTAATAGAC TGGATGGAGG CGGATAAAAGT
 4561 TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG
 4621 AGCCGGTGAG CGTGGGTCTC CGGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC
 4681 CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA
 4741 GATCGCTGAG ATAGGTGCCT CACTGATTA GCATTGGTAA CTGTCAGACC AAGTTACTC
 4801 ATATATACCT TAGATTGATT TAAAACCTCA TTTTTAATTT AAAAGGATCT AGGTGAAGAT
 4861 CCTTTTTGAT ATCTCATGA CAAAAATCCC TTAACGTGAG TTTTCGTTCC ACTGAGCGTC
 4921 AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCTCT TTTTTCTGC GCGTAATCTG
 4981 CTGCTTGCAA ACAAACACAC CACCGCTACC AGCGGTGGTT TGTGGCCGG ATCAAGAGCT
 5041 ACCAACTCTT TTTCCGAGG TAATCTGGTT CAGCAGAGCG CAGATACCAA ATACTGTCCT
 5101 TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT GTAGCACCGC CTACATACCT
 5161 CGCTCTGCTA ATCTGTAC CAGTGGCTGC TGCCAGTGGC GATAAGTCGT GTCTTACCGG
 5221 GTTGGACTCA AGACGATAGT TACCGATAA GGCGCAGCGG TCGGGCTGAA CGGGGGGTTC
 5281 GTGACACACAG CCCAGCTGG AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA
 5341 GCTATGAGAA AGGCCACGC TTCCCGAAGG GAGAAAGGCG GACAGGTATC CGGTAAGCGG
 5401 CAGGGTCGGA ACAGGAGAGC GCACGAGGG ACGTCCAGGG GGAAACGCCT GGTATCTTTA
 5461 TAGTCTGTC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA TTTTTGTGAT GCTCGTCAGG
 5521 GGGCGGAGC CTATGGAAAA ACGCCAGCA CGCGGCCCTT TTACGGTTCC TGGCCTTTG
 5581 CTGGCCCTTT GCTCACATGT TCTTCTGCG TTATCCCT GATTCTGTGG ATAACCGTAT
 5641 TACCGCCCTT GAGTGTGAGCTG ATACCGCTCG CGCGAGCCGA ACGACCGAGC GCAGCGAGTC
 5701 AGTGAGCGAG GAAGCGGAAG AGCGCTGAT GCGGTATTTT CTCCCTACGC ATCTGTGCGG
 5761 TATTCACAC CGCATATAATT TGTTAAAATT CGCGTTAAAT TTTGTAAAT TCAGCTCATT
 5821 TTTTAACCAA TAGGCCGAAA TCGGCAAAAT CCCTTATAAA TCAAAGAAT AGACCGAGAT
 5881 AGGGTTGAGT GTTGTGAGT TTGGAACAA GAGTCCACTA TTAAAGAAGC TGGACTCCAA
 5941 CGTCAAAGGG CGAAAAACCG TCTATCAGGG CGATGGCCCA CTACGTGAAC CATCACCTA
 6001 ATCAAGTTT TTGGGTGCA GGTGCCGTA AGCAGTAAAT CGGAACCCCTA AAGGGAGCCC-

FIGURE 24C

47/240

6061 CCGATTTAGA GCTTGACGGG GAAAGCCGGC GAACGTGGCG AGAAAGGAAG GGAAGAAAGC
6121 GAAAGGAGCG GGGCGTAGGG CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG TAACCAC
6181 ACCCGCCGCG CTTAATGCGC CGCTACAGGG CGCGTCCATT CGCCATTCAAG GCTGCTATGG
6241 TGCACTCTCA GTACAATCTG CTCTGATGCC GCATAGTTAA GCCAGTATAAC ACTCCGCTAT
6301 CGCTACGTGA CTGGGTCATG GCTGCGCCCC GACACCCGCC AACACCCGCT GACGCCCT
6361 GACGGGCTTG TCTGCTCCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT
6421 GCATGTGTCA GAGGTTTTCA CCGTCATCAC CGAACGCGC GAGGCAGCAG ATCAATTGCG
6481 GCGCGAAGGC GAAAGCGGCAT GCATTTACGT TGACACCATC GAATGGTGCA AAACCTTCG
6541 CGGTATGGCA TGATAGCGCC CGGAAGAGAG TCAATTCAAGG GTGGTGAATG TGAAACCAAGT
6601 AACGTTTATAC GATGTCGCAG AGTATGCCGG TGTCTCTTAT CAGACCGTTT CCCGCGTGGT
6661 GAACCAGGCC AGCCACGTTT CTGCGAAAAC GCGGGAAAAA GTGGAAGCGG CGATGGCGGA
6721 GCTGAATTAC ATTCCCAACC GCGTGGCACA ACAACTGGCG GGCAAACAGT CGTTGCTGAT
6781 TGGCGTTGCC ACCTCCAGTC TGGCCCTGCA CGCGCCGTG CAAATTGTCG CGGCATTAA
6841 ATCTCGCGCC GATCAACTGG GTGCCAGCGT GGTGGTGTG ATGGTAGAAC GAAGCGCGT
6901 CGAAGCCTGT AAAGCGCGG TGCACAAATCT TCTCGCGCAA CGCGTCAGTN GGGCTGATCA
6961 TTAA

FIGURE 24D

48/240

Figure 25A PDEST5

**pSPORT '+' (for sequencing, probes,
phagemid)**

1. agg cac ccc agg cct tac act tta tgc ttc egg ctc gaa tgt tgt gtg gaa
 tcc gtg ggg tcc gaa atg tga aat acg aag gcc gag cat aca gca cac ctt

52 reverse sequencing primers
ttg tga gct aac aat ttc aca cag gaa aca gct $\xrightarrow{\text{ATG}}$ "peptide
aac act cgc cta ttg tta aag tgt gtc ctt tgt cga tac tgg tac taa aeg
tgc

103 cca agc tct aat acg act cac tat agg gaa agc tgg tac gcc tgc atg cgg acg tcc tga gtg ata tcc gtt tcg acc atg cggt ^{Pst} ^{Kpn}

154 E_{co}KI Sma I Sal I ~~Int~~ ~~HRI~~
 egg tcc gta att ccc ggg tcg acg atc aca agt ttt zac zaa aaa act gaa
 gcc agg ect taa ggg ccc agc tgc tag tgt tca aac atg ttt tct cgt ott

↓ Gene

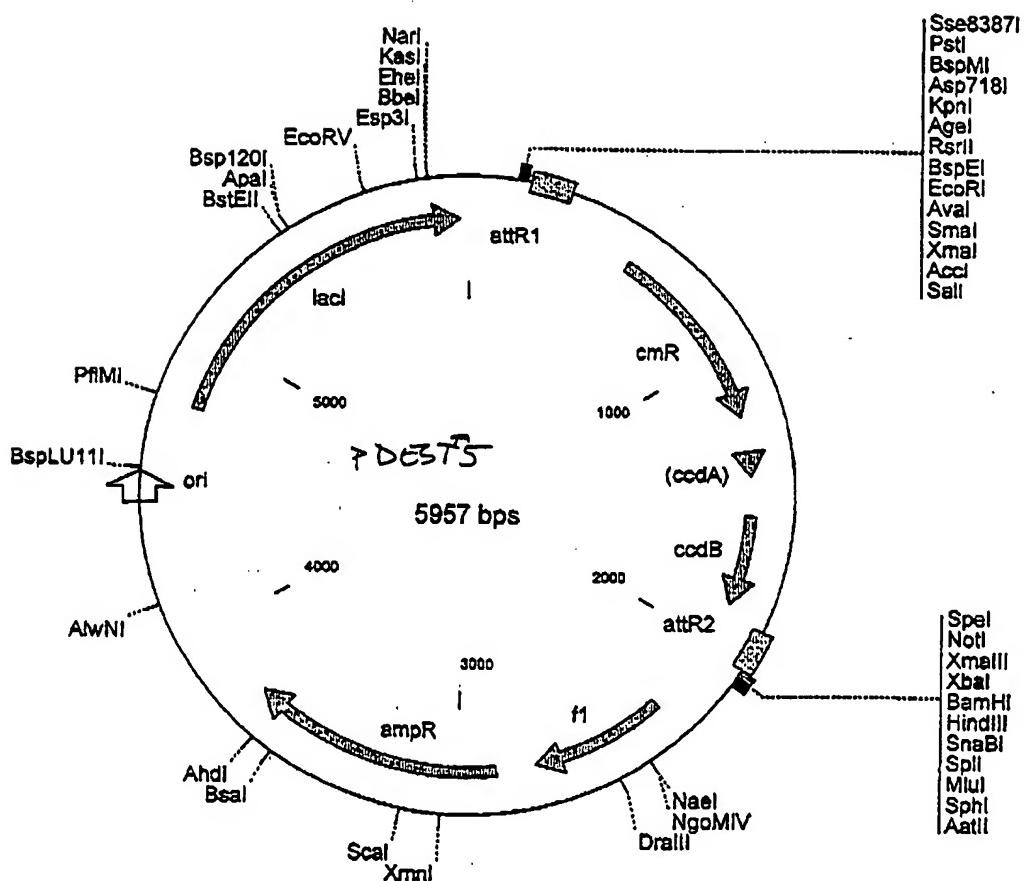
1990 Int ~~atR2~~ Spe
~~ttt acg ttt ctc gtt caa ctt det tgt aca aag tgg tga tca lcta gtc ggc~~
~~aaa tgc aaa gag caa gcc gac tga aca tgt ttc acc act agt gat cag ccg~~

2041 Not Xba Bam Hind 3 Mlu Sph
bgc cgc tct aga gga tcc atg cct acg tac acc tgc atg cga cgt cat agc
ccg gcg aga tct cct agg tcc gaa tgc atg cgc acg tac gct gca gta tcg

2092 tct tct aca gcg cca ccc aaa ttc aat tca ctg gcc gtc gtt tta caa cgt
aga aga tat cac agt gga ctt gag tta agt gac cgg cag caa aat gtt gca
SP6 Promoter
↓
SP6 RNA

2143 cg t gac tgg gaa aac cct ggc gtt acc caa ctt aat cgc ctt gca gca cat
gca ctg acc ctt ttg gga ccg caa tgg gtt gaa tta gcg gaa cgt cgt gta
. primers

Figure 25B *? DEST5* (cont'd)



50/240

pDEST5 5957 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
305..181	attR1
555..1214	CmR
1334..1418	inactivated ccdA
1556..1861	ccdB
1902..2026	attR2
2278..2733	f1 (f1 intergenic region)
2865..3722	ampR
5378..5538	ori
4756..5922	lacI

1 AGGCACCCCA GGCTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG
 61 GATAACAATT TCACACAGGA AACAGCTATG ACCATGATTA CGCCAAGCTC TAATACGACT
 121 CACTATAGGG AAAGCTGGTA CGCCTGCAGG TACCGGTCCG GAATTCCCGG GTCGACGATC
 181 ACAAGTTTGT ACAAAAAAGC TGAAACGAGAA ACGTAAAATG ATATAAATAT CAATATATTA
 241 AATTAGATTT TGCATAAAAA ACAGACTACA TAATACTGTA AAACACAACA TATCCAGTCA
 301 CTATGGCGGC CGCTAAGTTG GCAGCATCAC CGGACGCACT TTGCGCCGAA TAAATACCTG
 361 TGACGGAAGA TCACTTCGCA GAATAAATAA ATCCTGGTGT CCCTGTTGAT ACCGGGAAGC
 421 CCTGGGCCAA CTTTGGCGA AAATGAGACG TTGATCGGC CGTAAGAGGT TCCAACCTTC
 481 ACCATAATGA AATAAGATCA CTACCGGGCG TATTTTTGTA GTTATCGAGA TTTTCAGGAG
 541 CTAAGGAAGC TAAAATGGAG AAAAAAATCA CTGGATATAC CACCGTTGAT ATATCCAAT
 601 GGCATCGTAA AGAACATTT GAGGCATTT AGTCAGTTGC TCAATGTACC TATAACCAGA
 661 CCGTTCAGCT GGATATTAGC GCCTTTTAA AGACCGTAA GAAAAATAAG CACAAGTTT
 721 ATCCGGCTT TATTACATT CTTGCCCCC TGATGAATGC TCATCCGGAA TTCCGTATGG
 781 CAATGAAAGA CGGTGAGCTG GTGATATGGG ATAGTGTCA CCCTGTTAC ACCGTTTCC
 841 ATGAGCAAAC TGAAACGTT TCATCGCTC GGAGTGAATA CCACGACGAT TTCCGGCAGT
 901 TTCTACACAT ATATTCGAA GATGTGGCGT GTTACGGTGA AAACCTGGCC TATTCCTA
 961 AAGGGTTTAT TGAGAATATG TTTTCGGTCT CAGCCAATCC CTGGGTGAGT TTCACCAAGTT
 1021 TTGATTTAAA CGTGGCCAAT ATGGACAAC TCTTCGCCCC CGTTTTTCACC ATGGGCAAAT
 1081 ATTATACGCA AGGCGACAAG GTGCTGATGC CGCTGGCGAT TCAGGTTCAT CATGCCGTCT
 1141 GTGATGGCTT CCATGTCGGC AGAATGCTTA ATGAATTACA ACAGTACTGC GATGAGTGGC
 1201 AGGGCGGGC GTAAACGCGT GGATCCGGCT TACTAAAAGC CAGATAACAG TATGCGTATT
 1261 TGCGCGCTGA TTTTGCGGT ATAAGAATAT ATACTGATAT GTATACCCGA AGTATGTCAA
 1321 AAAGAGGTGT GCTATGAAGC AGCGTATTAC AGTGCAGTT GACAGCGACA GCTATCAGTT
 1381 GCTCAAGGCA TATATGATGT CAATATCTCC GGCTGGTAA GCACAACCAT GCAGAAATGAA
 1441 GCCCCGCGTC TCGTGGCGA ACGCTGGAAA GCGGAAAATC AGGAGGGAT GGCTGAGGTC
 1501 GCCCCGTTTA TTGAAATGAA CGGCTCTTT GCTGACGAGA ACAGGGACTG GTGAAATGCA
 1561 GTTTAAGGTT TACACCTATA AAAGAGAGAG CGGTTATCGT CTGTTGTGG ATGTACAGAG
 1621 TGATATTATT GACACGCCCG GGCGACGGAT GGTGATCCCC CTGGCCAGTG CACGTCCTGCT
 1681 GTCAGATAAA GTCTCCGTG AACTTACCC GGTGGTGAT ATCGGGGATG AAAGCTGGCG
 1741 CATGATGACC ACCGATATGG CCAGTGTGCC GGTCCTCGTT ATCGGGGAAG AAGTGGCTGA
 1801 TCTCAGCCAC CGCGAAAATG ACATCAAAAA CGCCATTAAC CTGATGTTCT GGGGAATATA
 1861 AATGTCAGGC TCCCTTATAC ACAGCCAGTC TGCAAGTCGA CCATAGTGAC TGGATATGTT
 1921 GTGTTTACA GTATTATGTA GTCTGTTTT TATGCAAAAT CTAATTAAAT ATATTGATAT
 1981 TTATATCATT TTACGTTCT CGITCAGCTT TCTTGTACAA AGTGGTGATC ACTAGTCGGC
 2041 GGCGCTCTA GAGGATCCAA GCTTACGTAC GCGTGCATGC GACGTCTAG CTCTTCTATA
 2101 GTGTCACCTA AATTCAATTG ACTGGCCGTC GTTTTACAAC GTCGTACTG GGAAAACCT
 2161 GGCGTTACCC AACTTAATCG CCTTGAGCA CATCCCCCTT TCGCCAGCTG GCGTAATAGC
 2221 GAAGAGGCC CGACCGATCG CCCTTCCCAA CAGTTGCGCA GCCTGAATGG CGAATGGACG
 2281 CGCCCTGTAG CGGCGCATTA AGCGCGCGG GTGTTGGTGT TACGCGCAGC GTGACCGCTA
 2341 CACTGCCAG CGCCCTAGCG CCCGCTCTT TCGCTTCTT CCCTTCTTT CTCGCCACGT
 2401 TCGCCGGCTT TCCCGTCAA GCTCTAAATC GGGGGCTCCC TTTAGGGTTC CGATTTAGTG
 2461 CTTTACGGCA CCTCGACCC AAAAACTTG ATTAGGGTGA TGTTTACGT AGTGGCCAT
 2521 CGCCCTGATA GACGGTTTT CGCCCTTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC
 2581 TCTTGTCTCA AACTGGAACA ACACCTCAACC CTATCTCGGT CTATTCTTT GATTATATAAG-

FIGURE 25C

51/240

2641 GGATTTGCC GATTCGGCC TATTGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG
 2701 CGAATTTAA CAAAATATTA ACCTTACAA TTTCAGGTGG CACTTTCGG GGAAATGTGC
 2761 GCGGAACCCC TATTTGTTA TTTTTCTAAA TACATTCAA TATGTATCCG CTCATGAGAC
 2821 AATAACCCCTG ATAAATGCTT CAATAATATT GAAAAAGGAA GAGTATGAGT ATTCAACATT
 2881 TCCGTGTCGC CCTTATTCCC TTTTTGCGG CATTTCGCT TCCCTGTTT GCTCACCCAG
 2941 AAACGCTGGT GAAAGTAAA GATGCTGAAG ATCAGTTGG TGACAGTG GGTTACATCG
 3001 AACTGGATCT CAACAGCGGT AAGATCCTTG AGAGTTTCG CCCCGAAGAA CGTTTCCAA
 3061 TGATGAGCAC TTTAAAGTT CTGCTATGTG GCGCGGTATT ATCCCGTATT GACGCCGGC
 3121 AAGAGCAACT CGGTGCGCGC ATACACTATT CTCAGAATGA CTTGGTTGAG TACTCACCAG
 3181 TCACAGAAAA GCATCTTAAG GATGGCATGA CAGTAAGAGA ATTATGCAGT GCTGCCATAA
 3241 CCATGAGTGA TAACACTGCG GCCAACCTAC TTCTGACAAC GATCGGAGGA CCGAAGGAGC
 3301 TAACCGCTTT TTTGACAAAC ATGGGGGATC ATGTAACCTG CCTTGATCGT TGGGAACCGG
 3361 AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCCAC GATGCCGTGA GCAATGGCAA
 3421 CAACGTTGCG CAAACTATTA ACTGGCGAAC TACTTACTCT AGCTTCCCGG CAACAATTAA
 3481 TAGACTGGAT GGAGGCGGAT AAAGITGCAG GACCACTCTC GCGCTCGGG CTTCCGGCTG
 3541 GCTGGTTTAT TGCTGATAAA TCTGGAGCGG GTGAGCGTGG GTCTCGGGT ATCATTGCAG
 3601 CACTGGGGCC AGATGGTAAG CCTCCCGTA TCGTAGTTAT CTACACGAGC GGGAGTCAGG
 3661 CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG TGCCCTACTG ATTAAGCATT
 3721 GTTAACGTC AGACCAAGTT TACTCATATA TACTTTAGAT TGATTTAAA CTTCATTTT
 3781 AATTAAAAG GATCTAGGTG AAGATCCTTT TTGATAATCT CATGACCAAA ATCCCTAAC
 3841 GTGAGTTTC GTTCCACTGA GCGTCAGACC CCGTAGAAA GATCAAAGGA TCTTCTGAG
 3901 ATCCTTTTT TCTGCGCGTA ATCTGCTGCT TGCAAACAAA AAAACCAACCG CTACCAAGCGG
 3961 TGGTTTGTGTT GCGGATCAA GAGCTACCAA CTCTTTTCC GAAGGTAACG GGCTTCAGCA
 4021 GAGCGCAGAT ACCAAATACT GTCCCTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA
 4081 ACTCTGTAGC ACCGCTACA TACCTCGCTC TGCTAATCCT GTTACCACTG GCTGCTGCCA
 4141 GTGGCGATAA GTCCGTGCTT ACCGGGGTGG ACTCAAGACG ATAGTTACCG GATAAGGCGC
 4201 AGCGGTCGGG CTGAACGGGG GGTTCTGCA CACAGCCCAG CTGGAGCGA ACGACCTACA
 4261 CCGAACTGAG ATACCTACAG CGTGAGCATT GAGAAAGCGC CACGCTTCCC GAAGGGAGAA
 4321 AGGCGGACAG GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCAGC AGGGAGCTTC
 4381 CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTCGGGTT TCGCCACCTC TGACTTGAGC
 4441 GTCGATTTTT GTGATGCTCG TCAGGGGGC GGAGCCTATG GAAAAACGCC AGCAACGCGG
 4501 CCTTTTTACG GTTCCCTGGCC TTTTGTGCGC CTCTTGCTCA CATGTTCTTT CCTGCGTTAT
 4561 CCCCTGATTC TGTGGATAAC CGTATTACCG CCTTTGAGTG AGCTGATAAC GCTCCCGCA
 4621 GCCGAACGAC CGAGCGCAGC GAGTCAGTGA GCGAGGAAGC GGAAGAGCGC CCAATACGCA
 4681 AACCGCCTCT CCCCCGCGGT TGGCGGATTG ATTAATGCAG AGCTTGCAAT TCGCCCGCGA
 4741 AGGCGAAGCG GCATTTACGT TGACACCATC GAATGGCGCA AAACCTTTCG CGGTATGGCA
 4801 TGATAGCGCC CGGAAGAGAG TCAATTCAGG GTGGTGAATG TGAAACCACT AACGTTATAC
 4861 GATGTCGCAAG AGTATGCCG TGTCTCTTAT CAGACCGTTT CCGCGTGGT GAACCAGGCC
 4921 AGCCACGTTT CTGCGAAAC GCGGGAAAAA GTGGAAGCGG CGATGGCGGA GCTGAATTAC
 4981 ATTCCCAACC GCGTGGCACA ACAACTGGCG GGCAAAACAGT CGTTGCTGAT TGGCGTTGCC
 5041 ACCTCCAGTC TGGCCCTGCA CGCGCGCTCG CAAATTGTCG CGCGATTAA ATCTCGCGCC
 5101 GATCAACTGG GTGCCAGCGT GTGGTGTGCG ATGGTAAAC GAAGCGCGT CGAACGCTGT
 5161 AAAGCGGCGG TGCACAATCT TCTCGCGCA CGGGTCAGTG GGCTGATCAT TAACTATCCG
 5221 CTGGATGACC AGGATGCCAT TGCTGTGAA GCTGCCGTCA CTAATGTTCC GGCGTTATTT
 5281 CTTGATGTCT CTGACCAGAC ACCCATCAAC AGTATTATTT TCTCCCATGA AGACGGTACG
 5341 CGACTGGCG TGGAGCATCT GGTCCCATG GGTACCCAGC AAATCGCGCT GTTACGCGGC
 5401 CCATTAAGTT CTGTCGGC GCGTCGCGT CTGGCTGGCT GGCATAAATA TCTCACTCGC
 5461 AATCAAATTC AGCCGATAGC GGAACGGGA GGGCACTGGA GTGCCATGTC CGGTTTCAA
 5521 CAAACCATGC AAATGCTGA TGAGGGCATC GTTCCCACTG CGATGCTGGT TGCCAACGAT
 5581 CAGATGGCGC TGGCGCAAT CGCGCCATT ACCGAGTCCG GGCTGCGCGT TGGTGCAGGAT
 5641 ATCTCGGTAG TGGGATACGA CGATACCGAA GACAGCTCAT GTTATATCCC GCCGTCAACC
 5701 ACCATCAAAC AGGATTTTCG CCTGCTGGGG CAAACCAAGCG TGGACCGCTT GCTGCAACTC
 5761 TCTCAGGGCC AGGCGGTGAA GGGCAATCAG CTGTTGCCG TCTCACTGGT GAAAAGAAAA
 5821 ACCACCCCTGG CGCCCAATAC GCAAAACGCC TCTCCCCCGC CGTTGGCCGA TTCATTAATG
 5881 CAGCTGGCAC GACAGGTTTC CCGACTGGAA AGCGGGCACT GAGCGCAACG CAATTAATGT
 5941 GAGTTAGCTC ACTCATT

FIGURE 25D

52/240

Figure 26A

pDEST6

pSPORT " " (opposite strand)

"forward" sequencing primers

1 taa/cgc cag ggt ttt ccc agt cac gac gtt gta aaa cga cgg cca gtg aat
att gcg gtc cca aaa ggg tca gtg ctg caa cat ttt get gcc ggt cac tta

52 SP6 promoter Sph Mlu
tga att tag gtg aca cta tag aag agc tat gac gtc gca tgc tgt acg
act caa atc cac tgt gat atc ttc tcc ata ctg cag tgt acg tgc gca tgc

103 Hind3 Bam Xba Not Spe Kpn R1 Int
~~tct~~ ~~acc~~ ~~tgt~~ ~~atc~~ ~~cac~~ ~~tgc~~ ~~agc~~ ~~ggc~~ ~~egc~~ ~~cgat~~ ~~cta~~ ~~gtg~~ ~~atc~~ ~~aca~~ ~~tgt~~ ~~ggc~~
~~att~~ ~~cga~~ ~~acc~~ ~~tag~~ ~~gag~~ ~~atc~~ ~~tgc~~ ~~ccg~~ ~~get~~ ~~gat~~ ~~cac~~ ~~tag~~ ~~tgt~~ ~~tca~~ ~~aac~~ ~~atg~~

154 ~~aaa~~ ~~aaa~~ ~~get~~ ~~ggc~~ ~~cga~~ ~~gaa~~ ~~acc~~ ~~tat~~ ~~aat~~ ~~gar~~ ~~ata~~ ~~aat~~ ~~atc~~ ~~aaa~~ ~~ata~~ ~~tta~~ ~~aat~~
~~ttt~~ ~~tat~~ ~~cga~~ ~~ttt~~ ~~get~~ ~~ttt~~ ~~tgc~~ ~~att~~ ~~tta~~ ~~cta~~ ~~tat~~ ~~tca~~ ~~tag~~ ~~tta~~ ~~tat~~ ~~aat~~ ~~tca~~

↓ Gene

1939 Int attR2
tat tta tat tat ttt acg ttt ctc gtt tag ctt ~~att~~ ~~tgt~~ ~~aca~~ ~~aag~~ ~~tgg~~ ~~tga~~
atc ~~att~~ ~~atc~~ ~~gtt~~ ~~aaa~~ ~~ttc~~ ~~acc~~ ~~gag~~ ~~gaa~~ ~~gtc~~ ~~gaa~~ ~~aca~~ ~~tgt~~ ~~ttc~~ ~~acc~~ ~~att~~

1990 Sal Sph EcoRI Kpn Pst
tcg tcc acc ggg gaa ttc ccg acc ggt atc tgc ggg tgt acc acc ttt ccc
agc agc ggg ggc ttt agg gcc tgg gca tgg tcc gca tgg tcc aaa ggg
T7 RNA

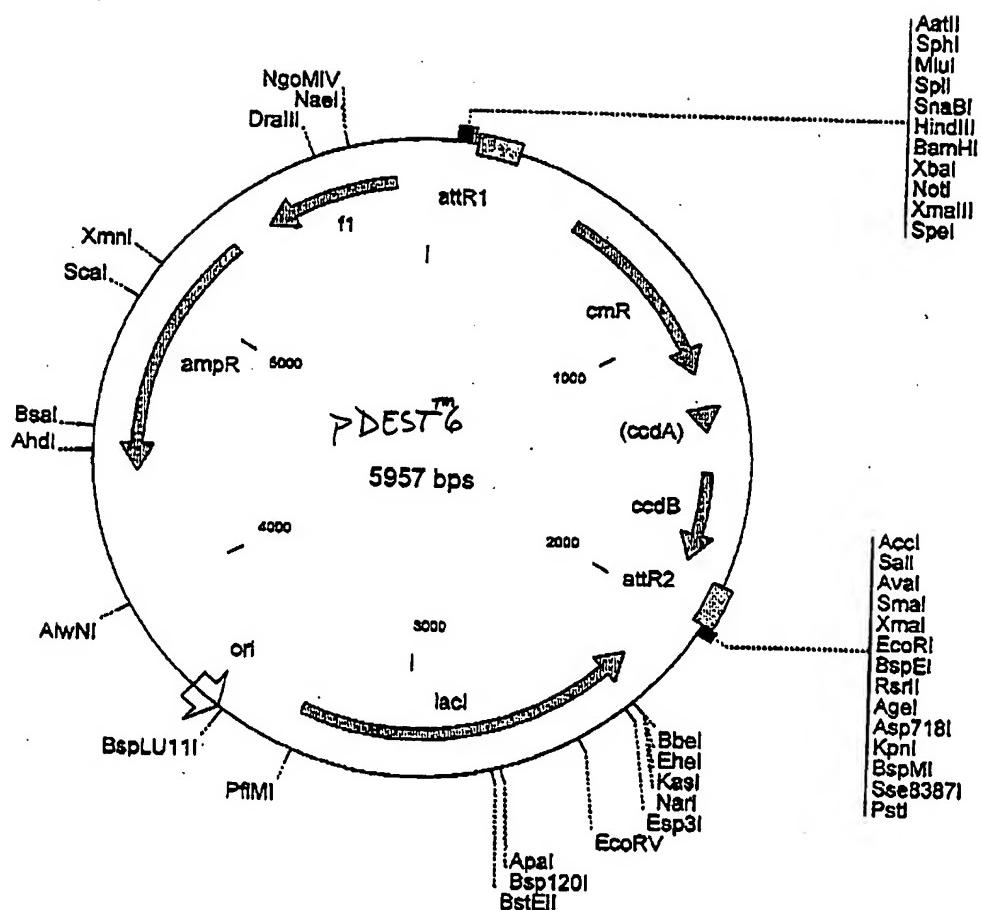
2041 tat agt gag tcc tat tag age ttg ggg taa tca tgg tca tag ctg ttt cct
ata tca ctc agc ata atc tcc aac cgc att agt acc tgt atc gac aaa gga
T7 promoter α-peptide "reverse .."

2092 gtg tga aat tgt tat ccg etc aca att cca cac ttt ata ⁻¹⁰ lac promoter
cac act tta aca ata ggc gag tgt taa ggt gtg ttt tat gtt ccg cct tcc
... sequencing primers lac RNA

2143 ata aag tgt aaa ggc tgg ggt ggc taa tga gtg age taa ctc aca tta att
tat ttc aca ttt ccg acc ccg att act cac tcc att gag tgt aat taa

S3/240

Figure 26B

PDEST6
(cont'd)

54/240

pDEST6 5957 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
266..142	attR1
516..1175	CmR
1295..1379	inactivated ccdA
1517..1822	ccdB
1863..1987	attR2
2203..3369	lacI
4403..5260	ampR
5392..5847	f1 (f1 intergenic region)

1 TAAACGCCAGG GTTTTCCCGAG TCACGCAGTT GTAAAACGAC GGCCAGTGAA TTGAATTAG
 61 GTGACACTAT AGAAGAGCTA TGACGTCGCA TGCACCGCGTA CGTAAGCTTG GATCCTCTAG
 121 AGCGGCCGCC GACTAGTGTG CACAAGTTTG TACAAAAAAAG CTGAACGAGA AACGTAAAAT
 181 GATAATAATA TCAATATATT AAATTAGATT TTGCATAAAA AACAGACTAC ATAATACTGT
 241 AAAACACAAAC ATATCCAGTC ACTATGGCGG CCCGCTAACGTT GGCAGCATCA CCCGACGCCAC
 301 TTTGCGCCGA ATAATAACCT GTGACGGAAG ATCACTTCGC AGAATAATA AATCCTGGTG
 361 TCCCTGTTGA TACCGGGAAAG CCCTGGGCCA ACTTTTGGCG AAAATGAGAC GTTGATCGGC
 421 ACGTAAGAGG TTCCAACCTT CACCATAATG AAATAAGATC ACTACCGGGC GTATTTTTG
 481 AGTTATCGAG ATTTTCAGGA GCTAAGGAAG CTAAAATGGA GAAAAAAATC ACTGGATATA
 541 CCACCGTTGA TATATCCCAA TGGCCTCGTA AAGAACATT TGAGGCATT TGAGTCAGTTG
 601 CTCATGTAC CTATAACCAAG ACCGTTCTAGC TGGATATTAC GGCCCTTTTA AAGACCGTAA
 661 AGAAAAATAA GCACAAGTTT TATCCGGCT TTATTCACAT TCTTGCCCGC CTGATGAATG
 721 CTCATCCCGA ATTCCGTATG GCAATGAAAG ACGGTGAGCT GGTGATATGG GATAGTGTTC
 781 ACCCTTGTAA CACCGTTTC CATGAGCAAA CTGAAACGTT TTCATCGCTC TGGAGTGAAT
 841 ACCACGACGA TTTCCGGCAG TTTCTACACA TATATTGCA AGATGTGGCG TGTACGGTG
 901 AAAACCTGGC CTATTTCCCT AAAGGGTTA TTGAGAATAT GTTTTCTGTC TCAGCCAATC
 961 CCTGGGTGAG TTTCACCAAGT TTTGATTAA ACGTGGCCAA TATGGACAAAC TTCTTCGCC
 1021 CCGTTTTCAC CATGGCAAA TATTATACGC AAGGCGACAA GGTGCTGATG CCGCTGGCGA
 1081 TTCAGGTTCA TCATGCGTC TGTGATGGCT TCCATGTCGG CAGAATGCTT AATGAATTAC
 1141 AACAGTACTG CGATGAGTGG CAGGGCGGGG CGTAAACGCG TGGATCCGGC TTACTAAAAG
 1201 CCAGATAAACCA GTATGCGTAT TTGCGCGCTG ATTGTTGCGG TATAAGAATA TATACTGATA
 1261 TGTATACCCG AAGTATGTCA AAAAGAGGTG TGCTATGAG CAGCGTATTA CAGTGACAGT
 1321 TGACAGCGAC AGCTATCAGT TGCTCAAGGC ATATATGATG TCAATATCTC CGGTCTGGTA
 1381 AGCACAAACCA TGCAGAAATGA AGCCCCGCGT CTGGCTGGCG AACGCTGGAA AGCGGAAAAT
 1441 CAGGAAGGGGAG TGGCTGAGGT ATTGAAATGA ACGGCTCTTT TGCTGACGAG
 1501 AACAGGGACT GTGAAATGC AGTTTAAGGT TTACACCTT AAAAGAGAGA GCCGTTATCG
 1561 TCTGTTTGTG GATGTACAGA GTGATATTAT TGACACGCC GGGCGACGGA TGGTGTACCCC
 1621 CCTGGCCAGT GCACGCTGTC TGTGAGATAA AGTCTCCCGT GAACTTTACC CGGTGGTGCA
 1681 TATCGGGGAT GAAAGCTGGC GCATGATGAC CACCGATATG GCCAGTGTGC CGGTCTCCGT
 1741 TATCGGGAA GAAAGTGGCTG ATCTCAGCCA CCGCGAAAAT GACATCAAAA AGCCATTAA
 1801 CCTGATGTTG TGGGGAAATAT AAAATGTCAGG CTCCCTTATA CACAGCCAGT CTGCAGGTG
 1861 ACCATAGTGA CTGGATATGT TGTGTTTAC AGTATTATGT AGTCTGTTT TTATGCAAAA
 1921 TCTAATTAA TATATTGATA TTTATATCAT TTTACGTTTC TCGTTCAGCT TTCTTGTACA
 1981 AAGTGGTGTAG CGTCGACCCCG GGAATTCCGG ACCGGTACCT GCAGGCGTAC CAGCTTCCC
 2041 TATAGTGTAGT CGTATTAGAG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT
 2101 TGTTATCCGC TCACAATTCC ACACAAACATA CGAGCCGGAA GCATAAAAGTG TAAAGCCTGG
 2161 GGTGCCATAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTCCCAG
 2221 TCGGGAAACC TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACGCGCGGG GAGAGGCGGT
 2281 TTGCGTATTG GGCAGCCAGGG TGGTTTTCT TTTCACCAAGT GAGACGGGCA ACAGCTGATT
 2341 GCCCTTCACC GCCTGGCCCT GAGAGAGTTG CAGCAAGCGG TCCACGCTGG TTTGCCCCAG
 2401 CAGCGAAAAA TCCTGTTGA TGGTGGTGA CGGCGGGATA TAACATGAGC TGTCTTGGT
 2461 ATCCGCGTAT CCCACTACCG AGATATCCGC ACCAACGCGC AGCCCGGACT CGGTAATGGC
 2521 GCGCATTGCG CCCAGCGCCA TCTGATCGTT GGCAACCAGC ATCGCAGTGG GAACGATGCC
 2581 CTCATTCAGC ATTTGATGG TTTGTTGAAA ACCGGACATG GCACCTCCAGT CGCCTTCCC
 2641 TTCCGCTATC GGCTGAATTG GATGCGAGT GAGATATTG TGCCAGGCCAG CCAGACCGAG-

FIGURE 26C

2701 ACGCGCCGAG ACAGAACTTA ATGGGCCGC TAACAGCGCG ATTGCTGGT GACCCATGC
 2761 GACCAGATGC TCCACGCCA GTCGCGTACG GTCTTCATGG GAGAAAATAA TACTGTGAT
 2821 GGGTGTCTGG TCAGAGACAT CAAGAAATAA CGCCCGAACAA TTAGTGCAGG CAGCTTCCAC
 2881 AGCAATGGCA TCCTGGTCAT CCAGCGGATA GTTAATGATC AGCCCACGTG CCCGTTGCGC
 2941 GAGAAGATTG TGCAACCGCCG CTTTACAGGC TTGACGCGCC CTCGGTCTA CCATCGACAC
 3001 CACCAACGCTG GCACCCAGTT GATCGGGCGC AGATTTAATC GCGCGACAA TTGCGACGG
 3061 CGCGTGCAGG GCCAGACTGG AGGTGGCACAC GCGAACATCAG AACGACTGTT TGCCCGCCAG
 3121 TTGTTGTGCC ACAGCGGTGG GAATGTAAATT CAGCTCCGCC ATCGCGCTT CCACCTTTTC
 3181 CGCGCTTTC CGAGAACGTT GGCTGGCTG GTTCACCCAGC CGGGAAACGG TCTGATAAGA
 3241 GACACCGCA TACTCTGCAG CATCGTATAA CGTTACTGGT TTACACATTCA CCACCCCTGAA
 3301 TTGACTCTCT TCCGGGGCCT ATCATGCCAT ACCCGCAGAAG GTTTTGGGCC ATTGATGGT
 3361 GTCAACGTAA ATGCCGCTTC GCCCTCGCGC GCGAACATTGA AGCTCTGCAT TAATGAATCG
 3421 GCGAACCGCG GGGGAGAGGC GGTTTGCCTA TTGGCGCTC TTCCGCTTCC TCGCTCACTG
 3481 ACTCGCTGCG CTCGGTCTGGT CGGCTCGGC GAGCGGTATC AGCTCACTCA AAGGCGGTAA
 3541 TACGGTTATC CACAGAAATCA GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC
 3601 AAAAGGCCAG GAACCGTAAA AAGGCGCGT TGCTGGCGTT TTTCCATAGG CTCCGGCCCC
 3661 CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAAGAGGTG GCGAACACCG ACAGGACTAT
 3721 AAAGATACCA GGCCTTCCC CCTGGAAGCT CCCTCGTGC CGTCTCTGTT CCGACCCCTGC
 3781 CGCTTACCGG ATACCTGTCC GCCTTCTCC CTTCGGGAAAG CGTGGCGCTT TCTCAATGCT
 3841 CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGTTCGCTC CAAGCTGGC TGTGTGCACG
 3901 AACCCCCCGT TCAGGCCGAC CGCTCGCCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC
 3961 CGGTAAGACA CGACTTATCG CCACCGCAG CAGCCACTGG TAACAGGATT AGCAGAGCGA
 4021 GGTATGTAGG CGGTGCTACA GAGTTCTGA AGTGGTGGCC TAACTACGGC TACACTAGAA
 4081 GGACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAC CTTCGGAAA AGAGTTGGTA
 4141 GCTCTTGATC CGGCAAACAA ACCACCGCTG GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC
 4201 AGATTAACGCG CAGAAAAAAA GGATCTCAAG AAGATCTTT GATCTTTCT ACGGGGCTG
 4261 ACGCTCAGTG GAACGAAAAC TCACGTTAAG GGATTTGGT CATGAGATTA TCAAAAGGA
 4321 TCTTACCTTA GATCCTTTTA AATTTAAAT GAAGTTTTAA ATCAATCTAA AGTATATATG
 4381 AGTAAACCTTG GTCTGACAGT TACCAATGCT TAATCAGTGA GGCACTATC TCAGCGATCT
 4441 GTCTATTTCG TTCACTCCATA GTTGCCTGAC TCCCCGCTCGT GTAGATAACT ACGATAACGGG
 4501 AGGGCTTACCC ATCTGGCCCC AGTGTGCAA TGATACCGCG AGACCCACGC TCACCCGCTC
 4561 CAGCTTATC AGCAATAAAAC CAGCCAGCGG GAAGGGCCGA GCGCAGAAGT GGTCTGCAA
 4621 CTTTATCCGC CTCCATCCAG TCTATTAAATT GTTGCCTGGGA AGCTAGAGTA AGTAGTTCGC
 4681 CAGTTAATAG TTTGCCAAC GTTGTGCCA TTGCTACAGG CATCGTGGTG TCACGCTCGT
 4741 CGTTTGGTAT GGCTTCATTC AGCTCCGGTT CCCAACGATC AAGGCGAGTT ACATGATCCC
 4801 CCATGTTGTG CAAAAAAAGCG GTTAGCTCCT TCAGGTCTCC GATCGTTGTC AGAAGTAAGT
 4861 TGGCCGCACTG GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT ACTGTCATGC
 4921 CATCCGTAAG ATGCTTTCT GTGACTGGTG AGTACTCAAC CAAGTCATTC TGAGAATAGT
 4981 GTATGCGCG ACCGAGTTGC TCTTGCCCG CGTCAATACG GGATAATACC GCGCACATA
 5041 GCAGAACTTT AAAAGTGCTC ATCATTGGAA AACGTTCTTC GGGCGAAA CTCTCAAGGA
 5101 TCTTACCGCT GTTGGAGATCC AGTTGATGT AACCCACTCG TGCAACCAAC TGATCTTCAG
 5161 CATCTTTAC TTTCACCCAGC GTTCTGGGT GAGAAAAAC AGGAAGGCAA AATGCCGCAA
 5221 AAAAGGAAAT AAGGGCGACA CGGAAATGTT GAATACTCAT ACTCTTCCTT TTCAATATT
 5281 ATTGAAGCAT TTATCAGGGT TATTGTCATCA TGAGCGGATA CATAATTGAA TGATTTAGA
 5341 AAAATAAAACA AATAGGGGTT CGCGCACAT TTCCCCGAAA AGTGCACCT GAAATTGTA
 5401 ACGTTAATAT TTTGTAAAAA TTGCGCTTAA ATTGTTGTTA AATCAGCTCA TTTTTAAC
 5461 AATAGGCCGA AATCGCAAA ATCCCTATA AATCAAAAGA ATAGACCGAG ATAGGGTTGA
 5521 GTGTTGTTCC AGTTTGGAAC AAGAGTCCAC TATTAAGAA CGTGGACTCC AACGTCAAAG
 5581 GCGAAACAC CGTCTATCAG GCGATGGCC CACTACGTGA ACCATCACCC TAATCAAGTT
 5641 TTTGGGGTC GAGGTGCCGT AAAGCACTAA ATCGGAACCC TAAAGGGAGC CCCCCGATTAA
 5701 GAGCTTGACG GGGAAAGCCG GCGAACCGTGG CGAGAAAGGA AGGGAAGAAA GCGAAAGGAG
 5761 CGGGCGCTAG GGCCTGGCA AGTGTAGCGG TCACGCTGC CGTAACCACC ACACCCGCCG
 5821 CGCTTAATGC GCCGCTACAG GGCCTGCTCA TTGCCCCATTG AGGCTGCGCA ACTGTTGGGA
 5881 AGGGCGATCG GTGCGGGCCTT CTTCGCTATT ACGCCAGCTG GCGAAAGGGG GATGTGCTGC
 5941 AAGGCGATTA AGTTGGG

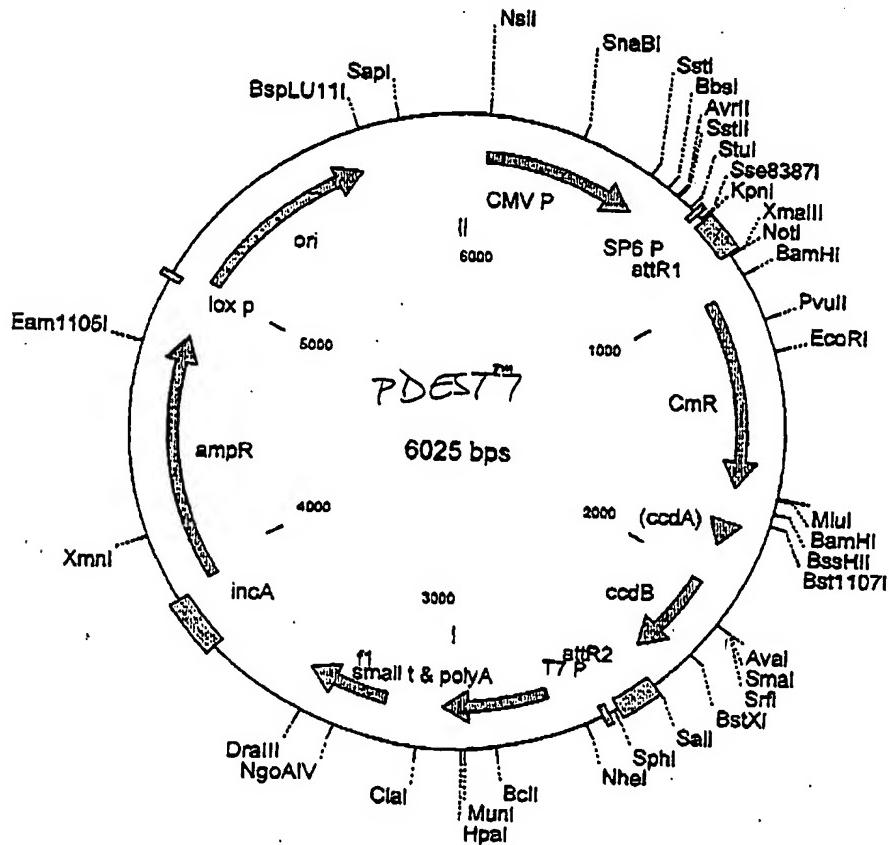
FIGURE 26d

56/240

Figure 27A: PDEST7

CMV promoter for eukaryotic expression

970 cca ttg acg caa atg ggc ggt agg cgt gta cgg tgg gag gtc tat ata agc
 ggt aac tgc gtt tac cgg cca tcc gca cat gcc acc ctc cag ata tat teg
 1021 aga gct cgt tta gtg aac cgt cag atc gcc tgg aga cgc cat cca cgc tgt
 tct cga gca aat cac ttg gca gtc tag cgg acc tct gcg gta ggt gcg aca
 CMV enhancer / promoter
 1072 ttt gac ctc cat aga aga cac cgg gac cga tcc agc ctc cgg act cta gca
 aaa ctg gag gta tct tct gtg gcc ctg gtc act agg teg gag gca tga gat cgg
 1123 tag gcc gcg gag cgg ata aca att tca cac agg aaa cag cta tga cca cta
 atc cgg cgc ctc gca tat tgt taa agt gtg tcc ttt gtc gat act ggt gat
 1174 ggc ttt tgc aaa aag cta ttt agg tga cac tat aga agg tac gcc tgc tgg
 ccg aaa acg ttt ttc gat aaa tcc act gtg ata tct tcc atg cgg acg tcc
 KpI EcoRI Int attR1
 1225 tac cgg tcc gga att ccc atc [aca agt ttg tag xaa xaa gat gaa cgg gaa
 atg gca agg cct taa ggg tag tgt tca aac atg ttt tct tca ctc oct etc]



571240

pDEST7 6025 bp (rotated to position 2800)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..589	CMV promoter
906..782	attR1
1015..1674	CmR
1794..1878	inactivated ccdA
2016..2321	ccdB
2362..2486	attR2
2671..3033	small t & polyA
3227..3502	f1
3962..4822	ampR
5022..5661	ori

1 ATTATCATGA CATTAAACCTA TAAAAATAGG CGTAGTACGA GGCCCTTTCA CTCATTAGAT
 61 GCATGTCGTT ACATAACTTA CGGTAAATGG CCCGCCTGGC TGACCGCCCA ACGACCCCCG
 121 CCCATTGACG TCAATAATGA CGTATGTTCC CATACTAACG CCAATAGGGG CTTTCCATTG
 181 ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCCTTG GCACTACATC AAGTGTATCA
 241 TATGCCAAGT ACGCCCCCTA TTGACGTCAA TGACGGAAA TGGCCCGCCT GGCATTATGC
 301 CCAGTACATG ACCTTATGGG ACTTTCTAC TTGGCACTAC ATCTACGTAT TAGTCATCGC
 361 TATTACCATG GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTGACTC
 421 ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTGTTT GGCACCAAAA
 481 TCAACGGGAC TTCCAAAAT GTCGTAACAA CTCCGCCCCA TTGACGCAA TGGCGGTAG
 541 GCGTGTACGG TGGGAGGTCT ATATAAGCAG AGCTCGTTA GTGAACCGTC AGATCGCTG
 601 GAGACGCCAT CCACGCTGTT TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG
 661 GACTCTAGCC TAGGCCGCGG AGCGGATAAC AATTCACAC AGGAAACAGC TATGACCATT
 721 AGGCCTTGC AAAAAGCTAT TTAGGTGACA CTATAGAAGG TACGCCCTGCA GGTACCGGAT
 781 CACAAGTTG TACAAAAAAAG CTGAACGGAGA AACGTTAAAT GATATAAATA TCAATATATT
 841 AAATTAGATT TTGCATAAAA AACAGACTAC ATAATACTGT AAAACACAAC ATATCCAGTC
 901 ACTATGGCGG CCGCATTAGG CACCCCAAGGC TTTCACCTTT ATGCTTCCGG CTCGTATAAT
 961 GTGTGGATTT TGAGTTAGGA TCCGTCGAGA TTTCAGGAG CTAAGGAAGC TAAAATGGAG
 1021 AAAAATCA CTGGATATAC CACCGTTGAT ATATCCAAT GGCATCGTAA AGAACATT
 1081 GAGGCATTTC AGTCAGTTG TCAATGTACC TATAACCGAG CGTTCAGCT GGATATTACG
 1141 GCCTTTAA AGACCGTAAA GAAAAAATAG CACAAGTTT ATCCGGCTT TATTACATT
 1201 CTTGGCCGCC TGATGAATGC TCATCCGGAA TTCCGTATGG CAATGAAAGA CGGTGAGCTG
 1261 GTGATATGGG ATAGTGTCA CCCTTGTAC ACCGTTTCC ATGAGCAAAC TGAAACGTT
 1321 TCATCGCTCT GGAGTGAATA CCACGACGAT TTCCGGCAGT TTCTACACAT ATATTGCAA
 1381 GATGTGGCGT GTTACGGGTAA AAACCTGGCC TATTTCCTA AAGGGTTTAT TGAGAATATG
 1441 TTTTCGTCT CAGCCAATCC CTGGGTGAGT TTCAACAGTT TTGATTTAAA CGTGGCCAAT
 1501 ATGGACAAC TCTTCGCCCC CGTTTTCACC ATGGGCAAT ATTATACGCA AGGCACAAAG
 1561 GTGCTGATGC CGCTGGCGAT TCAGGTTCAT CATGCCGTCT GTGATGGCTT CCATGTCGGC
 1621 AGAATGCTTA ATGAATTACA ACAGTACTGC GATGAGTGGC AGGGCGGGGC GTAAACCGGT
 1681 GGATCCGGCT TACTAAAAGC CAGATAACAG TATGCGTATT TGCGCGCTGA TTTTGCCTG
 1741 ATAAGAATAT ATACTGATAT GTATACCGA AGTATGTCAA AAAGAGGTGT GCTATGAAGC
 1801 AGCGTATTAC AGTGCACAGTT GACAGCGACA GCTATCAGTT GCTCAAGGCA TATATGATGT
 1861 CAATATCTCC GGTCTGGTAA GCACAAACCAT GCAGAATGAA GCCCGTCGTC TGCCTGCCGA
 1921 ACGCTGGAAA CGGGAAAATC AGGAAGGGAT GGCTGAGGTC GCCCGGTTA TTGAAATGAA
 1981 CGGCTCTTTT GCTGACGAGA ACAGGGACTG GTGAAATGCA GTTTAAGGTT TACACCTATA
 2041 AAAGAGAGAG CGGTTATCGT CTGTTTGTGG ATGTACAGAG TGATATTATT GACACGCCCG
 2101 GGCACGGAT GGTGATCCCC CTGGCCAGTG CACGCTGCT GTCAGATAAA GTCTCCGTG
 2161 AACTTACCC GGTGGTGCAT ATCGGGGATG AAAGCTGGCG CATGATGACCC ACCGATATGG
 2221 CCAGTGTGCC GGTCTCGTT ATCGGGGAAAG AAGTGGCTGA TCTCAGGCCAC CGCGAAAATG
 2281 ACATCAAAAA CGCCATTAAAC CTGATGTTCT GGGGAATATA AATGTCAGGC TCCCTTATAC
 2341 ACAGCCAGTC TGCAGGTGCA CCATAGTGAC TGGATATGTT GTGTTTACA GTATTATGTA
 2401 GTCTGTTTT TATGAAAAAT CTAATTTAAT ATATTGATAT TTATATCATT TTACGTTCT
 2461 CGTTCACTT TCTTGTACAA AGTGGTGATC GCGTGCATGC GACGTCATAG CTCTCTCCCT
 2521 ATAGTGAAGTC GTATTATAAG CTAGGCACTG GCCGTCGTT TACAACGTCG TGACTGGAA-

FIGURE 27B

88/240

2581 AACTGCTAGC TTGGGATCTT TGTGAAGGAA CCTTACTTCT GTGGTGTGAC ATAATTGGAC
 2641 AAACCTACCTA CAGAGATTTA AAGCTCTAAG GAAATATAA ATTTTTAAG TGTATAATGT
 2701 GTTAAACTAG CTGCATATGC TTGCTGCTTG AGAGTTTGCG TTACTGAGTA TGATTTATGA
 2761 AAATATTATA CACAGGAGCT AGTGAATTCA ATTGTTTG TGATTTAGAT TCACAGTCCC
 2821 AAGGCTCATT TCAGGGCCCT CAGTCCTCAC AGTCTGTTCA TGATCATAAT CAGCCATACC
 2881 ACATTGTAG AGGTTTACT TGCTTTAAA AACCTCCCAC ACCTCCCCCT GAACCTGAAA
 2941 CATAAAATGA ATGCAATTGT TGTTGTTAAC TTGTTTATTG CAGCTTATAA TGTTACAAA
 3001 TAAAGCAATA GCATCACAAA TTTCACAAAT AAAGCATTTT TTTCACTGCA TTCTAGTTGT
 3061 GGTTTGTCCA AACTCATCAA TGTTATCTTAT CATGCTGGA TGATCCTGC ATTATGAAAT
 3121 CGGCCAACGC GCGGGGAGAG GCGGTTTGC TATTGCTGG CGTAATAGCG AAGAGGCCG
 3181 CACCGATCGC CTTCCCAAC AGTTGCGCAG CCTGAATGGC GAATGGGACG CGCCCTGTAG
 3241 CGGCGCATT AAGCGCGGCG GTGTGGTGGT TACCGCAGC GTGACCGCTA CACTGCCAG
 3301 CGCCCTAGCG CCCGCTCCTT TCGCTTTCTT CCCCTCCTT CTGCCACGT TCGCCGGCTT
 3361 TCCCCGTCAA GCTCTAAATC GGGGGCTCCC TTAGGGTTC CGATTTAGTG CTTCACGGCA
 3421 CCTCGACCCC AAAAAGACTTG ATTAGGGTGA TGTTCACGT AGTGGGCCAT CGCCCTGTATA
 3481 GACGGTTTT CGCCCTTGTGA CGTTGGAGTC CACGTTCTT AATAGTGGAC TCTTGTCCA
 3541 AACTGGAACA ACACCTCAACC CTATCTCGGT CTATTCTTT GATTATAAG GGATTTGCC
 3601 GATTGGGCC TATTGGTTAA AAAATGAGCT GATTAAACAA AAATTTAACG CGAATTTAA
 3661 CAAAATATTA ACGTTTACAA TTTCAGGTGG CACTTTTCGG GGAAATGTGC GCGGAACCCC
 3721 TATTGTTA TTTTCTAA TACATTCAA TATGTATCCG CTATGCCAG GTCTTGACT
 3781 GGTGAGAACG GCTTGTCTGG CAGCTTCGAT GTGTGCTGGA GGGAGAATAA AGGTCTAAGA
 3841 TGTGCGATAG AGGGAAAGTCG CATTGAATTA TGTCGTGTT AGGGATCGCT GGTATCAAAT
 3901 ATGTGTGCC ACCCCTGGCA TGAGACAATA ACCCTGATAA ATGCTTCAT AATATTGAAA
 3961 AAGGAAGAGT ATGAGTATTAC AACATTCCG TGTCGCCCTT ATTCCCTTT TTGCGGCATT
 4021 TTGCTTCCT GTTTTGCTC ACCCAGAAAC GCTGGTGAAGA GTAAAAGATG CTGAAGATCA
 4081 GTTGGGTGCA CGAGTGGGTT ACATCGAAC GGATCTAAC AGCGGTAAAGA TCCCTTGAGAG
 4141 TTTTCGCCCC GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC
 4201 GGTATTATCC CGTATTGACG CGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA
 4261 GAATGACTTG GTTGAGTACT CACCACTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT
 4321 AAGAGAATTA TGCAGTGTCTG CCATAACCAT GAGTGATAAC ACTCGGGCCA ACTTACTTCT
 4381 GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTG CACAACATGG GGGATCATGT
 4441 AACTCGCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA
 4501 CACCAACGATG CCTGTAGCAA TGGCAACAAAC GTTGCCTCAA CTATTAACCTG GCGAACTACT
 4561 TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG CGGGATAAAG TTGCGAGGACC
 4621 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTATTGCT GATAAAATCTG GAGCCGGTGA
 4681 GCGTGGGTCT CGCGGTATCA TTGCACTGACT GGGGCCAGAT GGTAAAGCCCT CCCGTATCGT
 4741 AGTTATCTAC ACGACGGGGG GTCAGGCAAC TATGGATGAA CGAAAATAGAC AGATCGCTGA
 4801 GATAGGTGCC TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT
 4861 TTAGGATTGAT TAAAAACTTC ATTTTAATT TAAAAGGATC TAGGTGAAGA TCCCTTTGGA
 4921 TAATCTCATG CCATAACTTC GTATAATGTA TGCTATACGA AGTTATGGCA TGACCAAAAT
 4981 CCCTTAACGT GAGTTTCTG TCCACTGAGC GTCAAGACCCC GTAGAAAAGA TCAAAGGATC
 5041 TTCTTGAGAT CCTTTTTTC TGCGCTAAT CTGCTGCTTG CAAACAAAAA AACCAACCGCT
 5101 ACCAGCGGTG TTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTCCGA AGGTAACCTGG
 5161 CTTCAAGCAGA GCGCAGATAC CAAATACTGT CCTTCTAGTG TAGCCGTAGT TAGGCCACCA
 5221 CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCTGT TACCAAGTGGC
 5281 TGCTGCCAGT CGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA
 5341 TAAGGCCAG CGGTGGGGCT GAACGGGGG TTCGTGACAA CAGCCCCAGCT TGGAGCGAAC
 5401 GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCATTGA GAAAGCGCCA CGCTTCCCGA
 5461 AGGGAGAAAG CGGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG
 5521 GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCTT GTCGGGTTTC GCCACCTCTG
 5581 ACTTGAGCGT CGATTTTTGT GATGCTCGT AGGGGGGGGG AGCCATGGA AAAACCCAG
 5641 CAACCGGGCC TTTTACGGT TCCGGCCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTCC
 5701 TCGCTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGA CGTATACCGC
 5761 TCGCCGCAGC CGAACGACCG AGCGCAGCGA GTCACTGAGC GAGGAAGCGG AAGAGCGCCC
 5821 AATACGCAAAC CGCCCTCTCC CGCGCTGGTGC CGCGATTCA TAAATGCAAGAG CTTGCAATT
 5881 GCGCGTTTTT CAATATTATT GAAGCATTTA TCAGGGTTAT TGTCTCATGA GCGGATACAT
 5941 ATTTGAATGT ATTTAGAAAA ATAAACAAAT AGGGGTTCCG CGCACATTTC CCCGAAAAGT
 6001 GCCACCTGAC GTCTAAGAAA CCATT

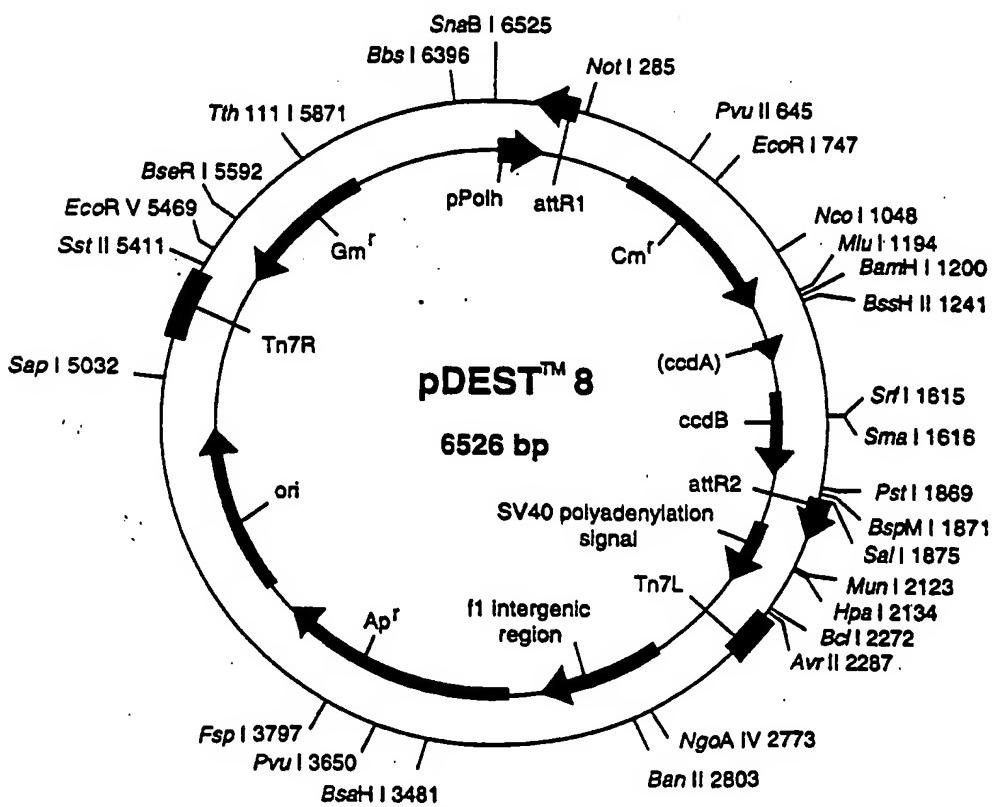
FIGURE 27C

59/240

Figure 78A: pDEST8 Polyhedron Promoter, Baculovirus Transfer Plasmid

AccI

1 cgt **ata** ctc cgg aat att aat aga tca tgg aga taa tta aaa tga taa cca
 gca **tat** gag gcc tta taa tta tct aat **acc** tct att aat ttt act att ggt
 ↓
 52 tet cgc aaa taa ata **agt** att tta'ctg ttt tcg taa cag ttt tgt aat aaa
 aga gcg ttt att tat tca taa aat gac aaa agc att gtc aaa aca tta ttt
 103 aaa acc tat aaa tat tcc gga tta ttc ata ccc tcc cac cat cgg ggg **agg**
 ttt tgg ata ttt ata agg cct aat aag tat ggc agg gtg gta gcc cgc gcc
 154 **(Bam)** **Int** **attR1** **attR2**
 atc atc **aca** **agt** **tgt** **tac** **aaa** **gtt** **gaa** **cgt** **gtt** **aa** **at** **at** **ata**
 tag tag **tgt** **tca** **aac** **atg** **ttt** **tcc** **cgt** **gtt** **ttt** **tgc** **att** **tta** **ctt** **tat**



60/240

pDEST8 6526 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
23..152	Ppolh
284..160	attR1
534..1193	CmR
1313..1397	inactivated ccda
1535..1840	ccdB
1881..2005	attR2
2766..3146	f1
3240..4090	ampR
4289..4869	ori
5564..6496	genR

1 CGTATACTCC GGAATATTAA TAGATCATGG AGATAATTAA AATGATAACC ATCTCGCAAA
 61 TAAATAAGTA TTTTACTGTT TTCGTAACAG TTTTGTAAATA AAAAACCTA TAAATATTCC
 121 GGATTATTCA TACCGTCCCCA CCATCGGGCG CGGATCATCA CAAGTTTGTA CAAAAAGCT
 181 GAACGAGAAA CGTAAATGA TATAAATATC AATATTAA ATTAGATTT GCATAAAAAAA
 241 CAGACTACAT AATACTGTAA AACACAACAT ATCCAGTCAC TATGGCGGCC GCTAAGTTGG
 301 CAGCATCACC CGACGCACTT TGCGCCGAAT AAATACCTGT GACGGAAGAT CACTTCGCAG
 361 AATAAATAAA TCCTGGTGT CCTGTTGATA CCGGGAAGCC CTGGGCCAAC TTTTGGCGAA
 421 AATGAGACGT TGATCGGCAC GTAAGAGGTT CCAACTTCA CCATAATGAA ATAAGATCAC
 481 TACCGGGCGT ATTTTTGAG TTATCGAGAT TTTCAGGAGC TAAGGAAGCT AAAATGGAGA
 541 AAAAACATCAC TGGATATACC ACCGTTGATA TATCCAATG GCATCGTAA GAACATTITG
 601 AGGCATTCA GTCAGTTGCT CAATGTACCT ATAACCAGAC CGTTCAGCTG GATATTACGG
 661 CCTTTTTAAA GACCGTAAAG AAAAATAAGC ACAAGTTTA TCCGGCTTT ATTACACATT
 721 TTGCCCCCCT GATGAATGCT CATCCGAAAT TCCGATGGC AATGAAAGAC GGTGAGCTGG
 781 TGATATGGGA TAGTGTTCAC CCTTGTACCA CCGTTTCCA TGAGCAAAC GAAACGTTT
 841 CATCGCTCTG GAGTGAATAC CACGACGATT TCCGGCAGTT TCTACACATA TATTGCAAG
 901 ATGTTGGCGTG TTACGGTGA AACCTGGCCT ATTCCCTAA AGGGTTTATT GAGAATATGT
 961 TTTTCGTCAGCCAAATCCC TGGGTGAGTT TCACCGAGTT TGATTTAAAC GTGGCCAATA
 1021 TGGACAACCTT CTTCGCCCCC GTTTTACCA TGGGCAAATA TTATACGCAA GGCGACAAGG
 1081 TGCTGATGCC GCTGGCGATT CAGGTTCATC ATGGCGTCTG TGATGGCTTC CATGTCGGCA
 1141 GAATGCTTAA TGAATTACAA CAGTACTGCG ATGAGTGGCA GGGCGGGGGC TAAACCGCTG
 1201 GATCCGGCTT ACTAAAAGCC AGATAACAGT ATGCGTATTG GCGCGCTGAT TTTTGGGT
 1261 TAAGAATATA TACTGATATG TATACCCGAA GTATGTCAAA AAGAGGTGTG CTATGAAGCA
 1321 GCGTATTACA GTGACAGTTG ACAGCGACAG CTATCAGTT CTCAAGGCAT ATATGATGTC
 1381 AATATCTCCG GTCTGGTAAG CACAACCATG CAGAATGAAG CCCGTGTCT GCGTGCAGAA
 1441 CGCTGGAAAG CGGAAAATCA GGAAGGGATG GCTGAGGTCG CCCGGTTTAT TGAAATGAAC
 1501 GGCTTTTG CTGACGAGAA CAGGGACTGG TGAAATGCG TTTAAGGTTT ACACCTATAA
 1561 AAGAGAGAGC CGTTATCGTC TGTTTGTGGA TGACAGAGT GATATTATTG ACACGCCGG
 1621 GCGACGGATG GTGATCCCCC TGGCCAGTGC ACGCTGCTG TCAGATAAAAG TCTCCCGTGA
 1681 ACTTTACCCG GTGGTGCATA TCGGGGATGA AAGCTGGCGC ATGATGACCA CCGATATGGC
 1741 CAGTGTGCCG GTCTCCGTTA TCGGGGAAGA AGTGGCTGAT CTCAGCCACC GCGAAAATGA
 1801 CATAAAAAAC GCCATTAACC TGATGTTCTG GGGAAATATAA ATGTCAGGCT CCCTTATACA
 1861 CAGCCAGTCT GCAGGTCGAC CATACTGACT GGATATGTT TGTTTACAG TATTATGTAG
 1921 TCTGTTTTT ATGAAAATC TAATTTAATA TATTGATATT TATATCATT TACGTTCTC
 1981 GTTCAGCTTT CTTGTACAA GTGGTGATAG CTTGTCGAGA AGTACTAGAG GATCATAATC
 2041 AGCCATACCA CATTGTAGA GGTTTTACTT GCTTTAAAAA ACCTCCCACA CCTCCCCCTG
 2101 AACCTGAAAC ATAAAATGAA TGCAATTGTT GTTGTAACT TGTTTATTGC AGCTTATAAT
 2161 GGTTACAAAT AAAGCAATAG CATCACAAAT TTCAACAAATA AAGCATTTTT TTCACTGCAT
 2221 TCTAGTTGTTG GTTTGTCCAA ACTCATCAAT GTATCTTATC ATGTCGGAT CTGATCACTG
 2281 CTTGAGCCTA GGAGATCCGA ACCAGATAAG TGAAATCTAG TTCCAAACTA TTTTGTCTT
 2341 TTTAATTTC GTATTAGCTT ACGACGCTAC ACCCAGTTCC CATCTATTG GTCACTCTTC
 2401 CCTAAATAAT CCTTAAAAAC TCCATTCCA CCCCTCCAG TTCCCAACTA TTTTGTCCGC
 2461 CCACAGCGGG GCATTTTCT TCCTGTTATG TTAAATCA AACATCCTGC CAACTCCATG
 2521 TGACAAACCG TCATCTTCGG CTACTTTTC TCTGTCACAG AATGAAAATT TTTCTGTCT

FIGURE 28B

61/240

2581 CTCTCGTTA TTAATGTTTG TAATTGACTG AATATCAACG CTTATTTGCA GCCTGAATGG
 2641 CGAATGGACG CGCCCTGTAG CGGCACATTG AGCCGGCGG GTGTGGTGGT TACGCAGC
 2701 GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCTT TCGCTTCTT CCCTTCCTT
 2761 CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC GGGGGCTCCC TTTAGGGTTC
 2821 CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAACCTTG ATTAGGGTGA TGGTTCACGT
 2881 AGTGGGCCAT CGCCCTGATA GACGGTTTT CGCCCTTGA CGTTGGAGTC CACGTTCTT
 2941 AATAGTGGAC TCTTGTCCA AACTGAAACA AACTCAACC CTATCTCGGT CTATTCTTT
 3001 GATTATAAG GGATTTGCC GATTCGGCC TATTGGTAA AAAATGAGCT GATTTAACAA
 3061 AAATTTAACG CGAATTTAA CAAAATATTA ACGTTACAA TTTCAGGTGG CACTTTCGG
 3121 GGAAATGTGC GCGGAACCCC TATTGTTA TTTTCTAA TACATTCAA TATGTATCCG
 3181 CTCATGAGAC ATAACCCCTG ATAAATGCTT CAATAATATT GAAAAGGAA GAGTATGAGT
 3241 ATTCAACATT TCCGTGTCGC CCTTATTCCC TTTTTGCGG CATTTCGCCT TCCTGTTTT
 3301 GCTCACCCAG AAACGCTGGT GAAAGTAAA GATGCTGAAG ATCAGTTGGG TGACCGAGTG
 3361 GGTTACATCG AACTGGATCT CAACAGCGGT AAGATCCTTG AGAGTTTCG CCCCAGAGAA
 3421 CGTTTCCAA TGATGAGCAC TTTTAAAGTT CTGCTATGTG CGCGGGTATT ATCCCCTATT
 3481 GACGCCGGGC AAGAGCAACT CGGTCGCCGC ATACACTATT CTCAGAATGA CTTGGTTGAG
 3541 TACTCACCAAG TCACAGAAAA GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGAGT
 3601 GCTGCCATAA CCATGAGTGA TAACACTGCG GCCAACCTAC TTCTGACAAAC GATCGGAGGA
 3661 CCGAAGGAGC TAACCGCTTT TTGACAAAC ATGGGGGATC ATGTAACCTG CCTTGATCGT
 3721 TGGGAACCCG AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCAAC GATGCCGTGA
 3781 GCAATGGCAA CAACCTTGGC CAAACTATTA ACTGGCGAAC TACTTACTCT AGCTTCCCG
 3841 CAACAATTAA TAGACTGGAT GGAGGCGGAT AAAGTTGCAG GACCACTTCT GCGCTCGGCC
 3901 CTTCCGGCTG GCTGGTTAT TGCTGATAAA TCTGGAGCCG GTGAGCGTGG GTCTCCGGT
 3961 ATCATTCGAG CACTGGGGCC AGATGGTAAG CCCTCCCGTA TCGTAGTTAT CTACACGACG
 4021 GGGAGTCAGG CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG TGCCTCACTG
 4081 ATTAAGCATT GGTAACTGTC AGACCAAGTT TACTCATATA TACTTTAGAT TGATTAAAA
 4141 CTTCATTTTT AATTAAAG GATCTAGGTG AAGATCCTTT TTGATAATCT CATGACCAAA
 4201 ATCCCTTAAC GTGAGTTTC GTTCCACTGA GCGTCAGACC CGTAGAAAA GATCAAAGGA
 4261 TCTTCTTGAG ATCCCTTTT TCTGCGCGTA ATCTGCTGCT TGCAAACAAA AAAACCAACG
 4321 CTACCAGCGG TGGTTGTTT GCGGATCAA GAGCTACCA CTCTTTTCC GAAGGTAACT
 4381 GGCTTCAGCA GAGCGCAGAT ACCAAATACT GTCTTCTAG TGTAGCCGTA GTTACCGCAC
 4441 CACTTCAAGA ACTCTGTAGC ACCGCTACAC TACCTCGCTC TGCTAATCT GTTACCGAGT
 4501 GCTGCTGCCA GTGGCGATAA GTCGTGTCTT ACCGGGTTGG ACTCAAGACG ATAGTTACCG
 4561 GATAAGGCAGC AGCGGTCGGG CTGACCGGG GGTCTGTCAC CACAGCCAG CTTGGAGCGA
 4621 ACGACCTACA CCGAACTGAG ATACCTACAG CGTGTAGCATT GAGAAAGCGC CACGCTTCCC
 4681 GAAGGGAGAA AGGCGGACAG GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCACG
 4741 AGGGAGCTTC CAGGGGGAAA CGCCCTGGTAT CTTTATAGTC CTGTCGGGTT TCGCCACCTC
 4801 TGACTTGAGC GTCGATTIT GTGATGCTCG TCAGGGGGGC GGAGCCTATG GAAAACGCC
 4861 AGCAACGCGG CCTTTTACG GTTCTGGCC TTTGCTGGC CTTTGCTCA CATGTTCTT
 4921 CCTGCGTTAT CCCCTGATTC TGTGGATAAC CGTATTACCG CTTTGAGTG AGCTGATACC
 4981 GCTGCCGCCA GCCAACGAC CGAGCGCAGC GAGTCAGTGA CACACCGCAG ACCAGCCGCG
 5041 CTGATGGGT ATTTCTCT TACGCATCTG TGCGGTATT TGCAAATTG CCCGCTGTAT
 5101 TAACTGGCA AAATCGGTTA CGGTTGAGTA ATAATGGAT GCCCTGCGTA AGCGGGTGTG
 5161 GGCGGACAAT AAGATCTTAA ACTGAACAAA ATAGATCTAA ACTATGACAA TAAAGTCTT
 5221 AACTAGACAG AATAGTTGTA AACTGAAATC AGTCCAGTTA TGCTGTAAA AAGCATACTG
 5281 GACTTTGTT ATGGCTAAAG CAAACTCTTC ATTTTCTGAA GTGCAAATTG CCCGCTGTAT
 5341 TAAAGAGGG CGTGGCCAAG GGCGATGGTAA AGACTATATT CGCGGCGTT TGACAATTAA
 5401 CCGAACAACT CGCGGGCCCG GAAGCCGATC TCGGCTTGAA CGAATTGTTA GGTGGCGGT
 5461 CTTGGGTCGA TATCAAAGTG CATCACTTCT TCCCGTATGC CCAACTTGT ATAGAGAGCC
 5521 ACTGCGGGAT CGTCACCGTA ATCTGCTTGC ACGTAGATCA CATAAGCACC AAGCGCGTTG
 5581 GCCTCATGCT TGAGGAGATT GATGAGCGCG GTGGCAATGC CCTGCCTCCG GTGCTCGCCG
 5641 GAGACTGCGA GATCATAGAT ATAGATCTCA CTACGCGGCT GCTCAAACCT GGGCAGAACG
 5701 TAAGCCGCCA GAGCGCCAAC AACCGCTCTT TGTCGAGG CAGCAAGCGC GATGAATGTC
 5761 TTACTACGGA GCAAGTTCCC GAGGTAAATCG GAGTCGGGCT GATGTTGGGA CTAGGTGGCT
 5821 ACGTCTCCGA ACTCACGACG GAAAAGATCA AGAGCAGGCC GCATGGATT GACTTGGTCA
 5881 GGGCCGAGCC TACATGTGGC AATGATGCCC ATACTTGAGC CACCTAACTT TGTTTAGGG
 5941 CGACTGCCCT GCTGCGTAAC ATCGTGTGCT CTGCGTAACA TCGTTGCTGC TCCATAACAT
 6001 CAAACATCGA CCCACGGCGT AACCGCGCTTG CTGCTTGGAT GCGCGAGGCA TAGACTGTAC-

FIGURE 28C

62/240

6061 AAAAAAACAG TCATAACAAG CCATGAAAAC CGCCACTGCG CCGTTACCAC CGCTGCCTTC
6121 GGTCAAGGTT CTGGACCAGT TGCCTGAGCG CATACTGCTAC TTGCATTACA GTTTACGAAC
6181 CGAACAGGCT TATGTCAACT GGGTTCTGTGC CTTCATCCGT TTCCACGGTG TGCGTCACCC
6241 GGCAACCTTG GGCAGCAGCG AAGTCGAGGC ATTTCTGTCC TGGCTGGCGA ACGAGCGCAA
6301 GGTTTCGGTC TCCACGCATC GTCAGGCATT GGCGGCCTTG CTGTTCTTCT ACGGCAAGGT
6361 GCTGTGCACG GATCTGCCCT GGCTTCAGGA GATCGGAAGA CCTCAGGCCGT CGCGGGCGCTT
6421 GCCGGTGGTG CTGACCCCCGG ATGAAGTGGT TCGCATCCTC GGTTTTCTGG AAGGGAGCA
6481 TCGTTTGTTC GCCCAGGACT CTAGCTATAG TTCTAGTGGT TGGCTA

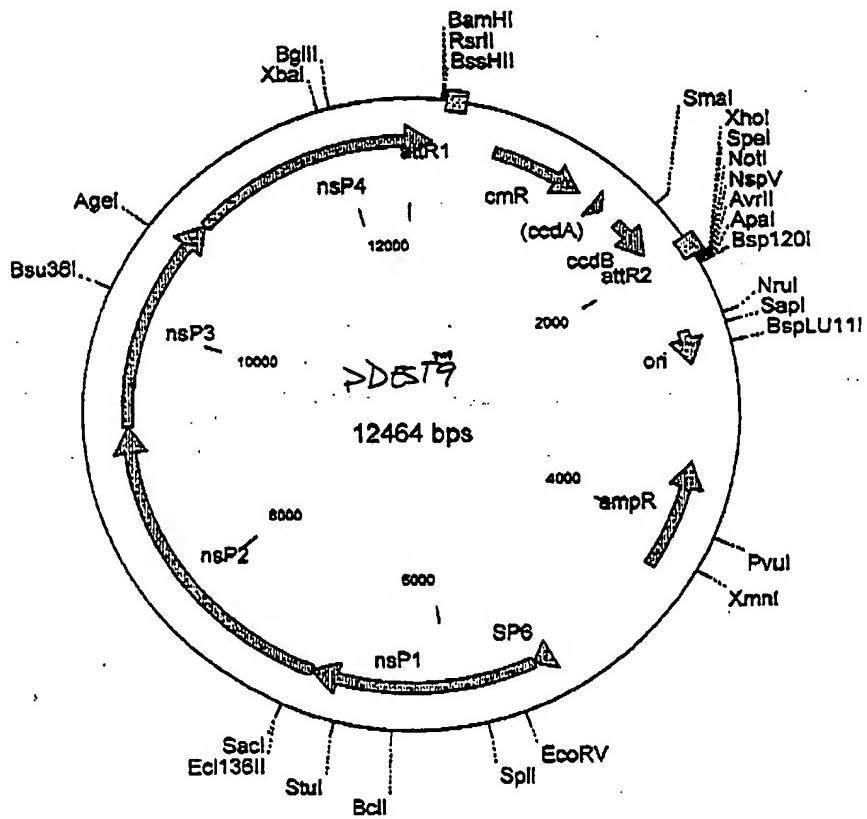
FIGURE 28D

63/240

Figure 29A: pDEST9

Semliki Forest Virus vector

103 ttg gcg agg gac att aag gcg ttt aag aaa ttg aga gga cct gtt ata ~~cgt~~
 aac cgc tcc ctg taa ttc cgc aaa ttc ttt aac tet cct gga caa tat ~~gtg~~
 154 ~~263 frames~~ → 265 RHA Bam
 ctc tac ggc ggt cct aca ~~ttt~~ ~~gtg~~ cgt taa tac aca gaa ttc tga ~~ttt~~ ~~gtt~~
 gag atg ~~ccg~~ ~~cca~~ gga tct aac ~~cac~~ ~~aca~~ gca att atg tgt ctt aag act aac cta
 205 Rsr II ~~Rsr I~~ R1
 ccc ~~agg~~ ~~ccg~~ aag cgc gct ttc cca tca ~~aca~~ ~~agt~~ ~~tgg~~ ~~tac~~ ~~aac~~ ~~aaa~~ ~~gtt~~ ~~gaa~~
 agg cca ~~ggc~~ ttc gcg cga aag ggt agt ~~tgt~~ ~~teo~~ ~~aac~~ ~~atg~~ ~~ttt~~ ~~ttt~~ ~~cgt~~ ~~ttt~~



pDEST9 12464 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
355..232	attR1
605..1264	CmR
1384..1468	inactivated ccda
1606..1911	ccdB
1952..2078	attR2
2532..2782	ori
3482..4282	ampR
5232..5365	SP6 promoter
5365..6965	nsP1:non-structural protein 1
6965..9265	nsP2:non-structural protein 2
9265..10865	nsP3:non-structural protein 3
10865..161	nsP4:non-structural protein 4

1 AGCAAGTGGT TCCGGACAGG CTTGGGGGCC GAACTGGAGG TGGCACTAAC ATCTAGGTAT
 61 GAGGTAGAGG GCTGCAAAAG TATCCTCATA GCCATGGCCA CCTTGGCGAG GGACATTAAG
 121 GCGTTTAAGA AATTGAGAGG ACCTGTTATA CACCTCTACG CGGGTCTAG ATTGGTGCCT
 181 TAATACACAG AATTCTGATT GGATCCCGGT CCGAAGCGCG CTTCCTCATC ACAAGTTTGT
 241 ACAAAAAAAGC TGAAACGAGAA ACGTAAAATG ATATAAAATAT CAATATATTA AATTAGATT
 301 TGCATAAAAAA ACAGACTACA TAATACTGTA AAACACAAACA TATCCAGTCA CTATGGCGGC
 361 CGCTAAGTTG GCAGCATCAC CCGACGCACT TTGCGCCGAA TAAATACCTG TGACGGAAGA
 421 TCACTTCGCA GAATAAAATAA ATCCTGGTGT CCCTGTTGAT ACCGGGAAGC CCTGGGCCAA
 481 CTTTTGGCGA AAATGAGACG TTGATCGCA CGTAAGAGGT TCCAACCTTC ACCATAATGA
 541 AATAAGATCA CTACCGGGCG TATTTTTGTA GTTATCGAGA TTTTCAGGAG CTAAGGAAGC
 601 TAAAATGGAG AAAAAAAATCA CTGGATATAC CACCGTTGAT ATATCCAAAT GGCACTCGTAA
 661 AGAACATTTC GAGGCATTTC AGTCAGTTGC TCAATGTACC TATAACCAGA CCGTTCAGCT
 721 GGATATTACG GCCTTTTAA AGACCGTAAA GAAAAATAAG CACAAGTTTT ATCCGGCCTT
 781 TATTACACATT CTTGCCCGCC TGATGAATGC TCATCCGGAA TTCCGTATGG CAATGAAAGA
 841 CGGTGAGCTG GTGATATGGG ATAGTGTCA CCCTGTTAC ACCGTTTTTC ATGAGCAAAC
 901 TGAAACGTTT TCATCGCTCT GGAGTGAATA CCACGACGAT TTCCGGCAGT TTCTACACAT
 961 ATATTGCAA GATGTGGCGT GTTACGGTGA AAACCTGGCC TATTTCCCTA AAGGGTTTAT
 1021 TGAGAATATG TTTTCGTCT CAGCCAATCC CTGGGTGAGT TTCACCAAGTT TTGATTAAAA
 1081 CGTGGCAAT ATGGACAATCT CTTCGCCCC CGTTTCACC ATGGGCAAAT ATTATACGCA
 1141 AGGCAGACAAG GTGCTGATGC CGCTGGCGAT TCAGGTTCAT CATGCCGTCT GTGATGGCTT
 1201 CCATGTCGGC AGAATGCTTA ATGAATTACA ACAGTACTGC GATGAGTGGC AGGGCGGGGC
 1261 GTAAAGATCT GGATCCGGCT TACTAAAAGC CAGATAACAG TATGCGTATT TGCAGCCTGA
 1321 TTTTGCGGT ATAAGAATAT ATACTGATAT GTATAACCGA AGTATGTCAA AAAGAGGTGT
 1381 GCTATGAAGC AGCGTATTAC AGTGACAGTT GACAGCGACA GCTATCAGTT GCTCAAGGCA
 1441 TATATGATGT CAATATCTCC GGTCTGGTAA GCACAAACCAT GCAGAATGAA GCCCCTCGTC
 1501 TCGTGGCGA ACGCTGGAAA GCGGAAAATC AGGAAGGGAT GGCTGAGGTC GCCCCTGGTTA
 1561 TTGAAATGAA CGGCTCTTT GCTGACGAGA ACAGGGACTG GTGAAATGCA GTTAAAGGTT
 1621 TACACCTATA AAAGAGAGAG CGGTTATCGT CTGTTTGTGG ATGTACAGAG TGATATTATT
 1681 GACACGCCCG GGCACGGAT GGTGATCCCC CTGGCCAGTG CACGCTCTGCT GTCAGATAAA
 1741 GTCTCCCGTG AACTTTACCC GGTGGTGCAT ATCGGGGATG AAAGCTGGCG CATGATGACC
 1801 ACCGATATGG CCAGTGTGCC GGTCTCCGGT ATCGGGGAG AAGTGGCTGA TCTCAGCCAC
 1861 CGCGAAAATG ACATCAAAA CGCCATTAAC CTGATGTTCT GGGGAATATA AATGTCAGGC
 1921 TCCCTTATAC ACAGCCAGTC TGCAGGTGCA CCATAGTGAC TGGATATGTT GTGTTTTACA
 1981 GTATTATGTA GTCTGTTTT TATGCAAAG TGCTAATTTA ATATATTGAT ATTTATATCA
 2041 TTTTACGTT CTCGTTCAAC AAACTGGTGA TGGGAACCTCG AGTTCACTAG
 2101 TCGATCCCGC GGCGCGCTTC GAACCTAGGC AAGCATGCGG GCCCAGTGGG TAATTAATTG
 2161 AATTACATCC CTACGCAAAC GTTTTACGGC CGCCCGTGGC GCCCAGCGCCC GGCGGGCCGT
 2221 CCTTGCCCGT TGCAGGCCAC TCCGGTGGCT CCCGCTCGTCC CCGACTTCCA GGCCCAGCAG
 2281 ATGCAGCAAC TCATCAGCGC CGTAAATGCG CTGACAATGA GACAGAACGC AATTGCTCCT
 2341 GCTAGGAGCT TAATTGACCG AATAATTGGA TTTTTATTTT ATTTTGCAAT TGGTTTTAA
 2401 TATTTCAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA-

Fig 1B 29B

65/240

2461 AAAAAAAA AAAAAAAACTA GAAATCGCGA TTTCTAGTCT GCATTAATG A ATCGGCCAAC
 2521 GCGCGGGGAG AGGCCGTTTG CGTATTGGC GCTCTCCGC TTCTCGCTC ACTGACTCGC
 2581 TGCCTCGGT CGTCGCGCTG CGCGAGCGG TATCAGCTCA CTCAAAGGC GTAATACGGT
 2641 TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG
 2701 CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTCTCA TAGGCTCCGC CCCCTGACG
 2761 AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT
 2821 ACCAGGCCTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCCGCTTA
 2881 CGGATACCT GTCCGCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA TGCTCGCGCT
 2941 GTAGGTATCT CAGTTCGGTG TAGGTCGTT GCTCCAAGCT GGGCTGTGTG CACGAACCCC
 3001 CGITTCAGCC CGACCGCTGC GCCTTATCCG GTAACTATCG TCTTGAGTCC AACCCGGTAA
 3061 GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG
 3121 TAGGCGGTGTC TACAGAGTT TTGAAGTGGT GGCTTAACCA CGGCTACACT AGAAGGACAG
 3181 TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT
 3241 GATCCGGCAA ACAAAACACC GCTGGTAGCG GTGGTTTTTG TGTTTGCAAG CAGCAGATT
 3301 CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC
 3361 AGTGGAACGA AAACCTCACGT TAAGGGATTT TGTCATGAG ATTATCAAAA AGGATCTTCA
 3421 CCTAGATCCT TTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA
 3481 CTTGGTCTGA CAGTTACCAA TGCTTAATCA TGAGGACCC TATCTCAGCG ATCTGCTAT
 3541 TTGTTTCATC CATAGTTGCC TGACTCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT
 3601 TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT
 3661 TATCAGCAAT AAACCAAGCCA GCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTAT
 3721 CGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAAGCTAG AGTAAGTAGT TCGCCAGTTA
 3781 ATAGTTTGC GCAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTG
 3841 GTATGGCTTC ATTCAAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT
 3901 TGTGCAAAAA AGCGGTTAGC TCCCTCGTC CTCCGATCGT TGTCAGAAGT AAGTTGCCG
 3961 CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCATAATT TCTTACTGTC ATGCCATCCG
 4021 TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC
 4081 GGCGACCGAG TTGCTCTTGC CGCGCTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA
 4141 CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGCG AAAACTCTCA AGGATCTTAC
 4201 CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT
 4261 TTACTTTCAC CAGCGTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC GCAAAAAGG
 4321 GAATAAGGC GACACGGAAA TGTTGAATAC TCATACTCTT CTTTTCTCAA TATTATTGAA
 4381 GCATTTATCA GGTTTATTGT CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA
 4441 AACAAATAGG GGTTCCGC ACATTTCCCC GAAAAGTGC ACCTGACGTC TAAGAAACCA
 4501 TTATTATCAT GACATTAACC TATAAAAATA GGCATGATCAC GAGGCCCTTT CGTCTCGCGC
 4561 GTTTGGTGA TGACGGTGAA AACCTCTGAC ACATGCAAGCT CCCGGAGACCG GTCACAGCTT
 4621 CTGCTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG CGCGTCAGCG GGTGTTGGCG
 4681 GGTGTCGGGG CTGGCTTAAC TATGCGGCAT CAGAGCAGAT TGTACTGAGA GTGCACCATA
 4741 TCGACCGCTCT CCCTTATGCC ACTCCTGCT TAGGAAAGCAG CCCAGTACTA GGTGAGGCC
 4801 GTTGAGCACC GCCGCCGCAA GGAATGGTGC ATGCAAGGAG ATGGCGCCCA ACAGTCCCC
 4861 GGCCACGGGG CCTGCCACCA TACCCACGCC GAAACAAGCG CTCATGAGCC CGAAGTGGCG
 4921 AGCCCCATCT TCCCCATCGG TGATGTCGGC GATATAGGCG CCAGCAACCG CACCTGTGGC
 4981 GCGGTGATG CCGGCCACGA TGCGTCCCGC GTAGAGGATC TGGCTAGCGA TGACCCCTGCT
 5041 GATTGGTTCG CTGACCATTT CCGGGGTGCG GAACGGCGTT ACCAGAAAAT CAGAAGGTTC
 5101 GTCCAACCAA ACCGACTCTG ACGGCAGTTT ACGAGAGAGA TGATAGGGTC TGCTTCAGTA
 5161 AGCCAGATGC TACACAATT GGCTTGACA TATTGTCGTT AGAACGCCG TACAATTAAAT
 5221 ACATAACCTT ATGTATCATA CACATACGAT TTAGGTGACA CTATAGATGG CGGATGTGTG
 5281 ACATACACGA CGCCAAAAGA TTTTGTCTCA GCTCCGCTAC CGAGAGAGATT
 5341 AACCACCCAC GATGGCCGCC AAAGTGCAATG TTGATATTGA GGCTGACAGC CCATTCAATCA
 5401 AGTCTTGC AAGGGCATT CCCTCGTTG AGGTGGAGTC ATTGCAGGTC ACACCAAATG
 5461 ACCATGCAAA TGCCAGAGCA TTTTCCGACCC TGGCTACCAA ATTGATCGAG CAGGAGACTG
 5521 ACAAAGACAC ACTCATCTTG GATATCGCA GTGCCCTTC CAGGAGAATG ATGTCTACGC
 5581 ACAAATACCA CTGCGTATGC CCTATCGCA CGCGAGAAGA CCCCCGAAAGG CTCGATAGCT
 5641 ACGCAAAGAA ACTGGCAGCG GCCTCCGGGA AGGTGCTGGA TAGAGAGATC GCAGGAAAAA
 5701 TCACCGACCT GCAGACCGTC ATGGCTACGC CAGACGCTGA ATCTCCTACC TTTTGCTGC
 5761 ATACAGACGT CACGTGTCGT ACGGCAGCCG AAGTGGCCGT ATACCAGGAC GTGTATGCTG
 5821 TACATGCACC AACATCGCTG TACCATCAGG CGATGAAAGG TGTCAGAACCG GCGTATTGGA
 5881 TTGGGTTTGA CACCACCCCG TTTATGTTG ACGCGCTAGC AGGCGCGTAT CCAACCTACG-

FIGURE 29C

66/240

5941 CCACAAACTG GGCGCAGCAG CAGGTGTTAC AGGCCAGGAA CATAGGACTG TGTGCAGCAT
 6001 CCTTGACTGA GGGAAAGACTC GGCAAACAGT CCATTCTCCG CAAGAAGCAA TTGAAACCTT
 6061 GCGACACAGT CATGTTCTCG GTAGGATCTA CATTGTACAC TGAGAGCAGA AAGCTACTGA
 6121 GGAGCTGGCA CTTACCCCTCC GTATTCCACC TGAAAGGTAA ACAATCCTTT ACCTGTAGGT
 6181 GCGATACCAT CGTATCATGT GAAGGGTAGC TAGTTAAGAA AATCACTATG TGCCCCGGCC
 6241 TGTACGGTAA AACGGTAGGG TACGGCGTGA CGTATCACCG GGAGGGATTG CTAGTGTGCA
 6301 AGACCACAGA CACTGTAAA GGAGAAAGAG TCTCATTCCC TGATGCACC TACGTCCCT
 6361 CAACCATCTG TGATCAAATG ACTGGCATAC TAGCGACCGA CGTCACACCG GAGGACGAC
 6421 AGAAGTTGTT AGTGGGATTG AATCAGAGGA TAGTTGTGAA CGGAAGAACAA CAGCGAAACA
 6481 CTAACACGAT GAAGAACTAT CTGCTTCCGA TTGTGGCCGT CGCATTAGC AAGTGGCGA
 6541 GGGAAATACAA GGCAGACCTT GATGATGAAA AACCTCTGGG TGCCGAGAG AGGTCACTTA
 6601 CTTGCTGCTG CTTGTGGCA TTAAAACGA GGAAGATGCA CACCATGTAC AAGAAACAG
 6661 ACACCCAGAC AATAGTGAAG GTGCCCTCAG AGTTAACCTC GTTCGTATC CCGAGCCTAT
 6721 GGTCTACAGG CCTCGCAATC CCAGTCAGAT CACCGATTAA GATGCTTTG GCCAAGAAGA
 6781 CCAAGCGAGA GTTAATACCT GTTCTCGACG CGTCGTCAAG CAGGGATGCT GAACAAGAGG
 6841 AGAAGGAGAG GTTGGAGGCC GAGCTGACTA GAGAACCTT ACCACCCCTC GTCCCCATCG
 6901 CGCCGGCGGA GACGGGAGTC GTGACGTG ACGTTGAAGA ACTAGAGTAT CACGCAGGTG
 6961 CAGGGGTCTGTT GGAAACACCT CGCAGCGCTG TGAAAGTCAC CGCACAGCCG AACGACGTAC
 7021 TACTAGGAAA TTACGTAGTT CTGCTCCCGC AGACCGTGT CAAGAGCTCC AAGTTGGCCC
 7081 CCGTGCACCC TCTAGCAGAG CAGGTGAAA TAATAACACA TAACGGGAGG GCCGGCGGTT
 7141 ACCAGGTGCA CGGATATGAC GGCAGGGTCC TACTACCATG TGATCGGCC ATTCCGGTCC
 7201 CTGAGTTCA GGCTTGAGC GAGAGCGCCA CTATGGTGTAA CAACGAAAGG GAGTTCTCA
 7261 ACAGGAAACT ATACCATATT GCGGTTACG GACCCCTCGCT GAACACCGAC GAGGAGAACT
 7321 ACGAGAAAGT CAGAGCTGAA AGAACTGACG CCGAGTACGT GTTCGACGTA GATAAAAAAT
 7381 GCTGCGTCAA GAGAGAGGAA GCGTCGGGTT TGGTGGTGGT GGGAGAGCTA ACCAACCCCC
 7441 CGTCCATGA ATTGCGCTAC GAAGGGCTGA AGATCAGGCC GTGGCACCATAAAGACTA
 7501 CAGTAGTAGG AGTCTTGGG GTTCCGGAT CAGGCAAGTC TGCTATTATT AAGAGCCTCG
 7561 TGACCAAACA CGATCTGGTC ACCAGCGGCA AGAAGGAGAA CTGCCAGGAA ATAGTTAACG
 7621 ACGTGAAGAA GCACCGCGGG AAGGGGACAA GTAGGGAAAA CAGTGAACCTCC ATCCCTGCTAA
 7681 ACGGGTGTCG TCGTGCCTGTG GACATCCTAT ATGTGGACGA GGCTTTCGCT TGCCATTCCG
 7741 GTACTCTGCT GGCCCTAATT GCTCTTGTAA AACCTCGGAG CAAAGTGGTG TTATGC3GAG
 7801 ACCCCAAGCA ATGCGGATTTC TTCAATATGA TGCACTTAA GGTGAACCTTC AACCACAAACA
 7861 TCTGCACTGA AGTATGTAT AAAAGTATAT CCAGACGTTG CACCGTCCA GTCACG3CCA
 7921 TCGTGTCTAC GTTGCCTAC GGAGGCAAGA TGCGCACGAC CAACCCGTGC AACAAACCCA
 7981 TAATCATAGA CACCACAGGA CAGACCAAGC CCAAGCCAGG AGACATCGT TTAACATGCT
 8041 TCCGAGGCTG GGCAAAGCAG CTGCACTTGG ACTACCGTGG ACACGAAGTC ATGACAGCAG
 8101 CAGCATCTCA GGGCCTCACC CGCAAAGGGG TATACCCGT AAGGCAGAAG GTGAAT3AAA
 8161 ATCCCTGTAA TGCCCTGCG TCAGGAGCACG TGAATGTACT GCTGACGCGC ACTGAGGATA
 8221 GGCTGGTGTG GAAAACGCTG GCCGGCGATC CCTGGATTAA GTTCTTATCA AACATTCCAC
 8281 AGGGTAACCT TACCGCCACA TTGGAAGAAT GGCAAGAAGA ACACGACAAA ATAATGAAGG
 8341 TGATTGAAGG ACCGGCTGCG CCTGTGGACG CGTTCCAGAA CAAAGCGAAC GTGTGTGGG
 8401 CGAAAAGCCT GGTGCGTGTG CTGGACACTG CGGAAATCG ATTGACAGCA GAGGAGTGG
 8461 GCACCCATAAT TACAGCATTT AAGGAGGACA GAGCTACTC TCCAGTGGTG GCCTTGAATG
 8521 AAATTTGCAC CAAGTACTAT GGAGTTGACC TGGACAGTGG CCTGTTTCT GCCCCGAAGG
 8581 TGCTCCGTAA TTACGAGAAC AACCACCTGGG ATAACAGACCC TGGTGGAAAGG ATGTAT3GAT
 8641 TCAATGCCGC AACAGCTGCC AGGCTGGAG CTAGACATAC CTTCCTGAAAG GGGCAGTGGC
 8701 ATACGGGCAA GCAGGAGCTT ATCCGAGAAA GAAAATCCA ACCGCTTCT GTGCTG3ACA
 8761 ATGTAATTCC TATCAACCGC AGGCTGCCGC AGCCCTGGT GGCTGAGTAC AAGACGGTTA
 8821 AAGGCAGTAG GTTGTGAGTGG CTGGTCAATA AAGTAAGAGG GTACCCACGTC CTGCTGGTGA
 8881 GTGAGTACAA CCTGGCTTGC CCTCGACGCC GGGTCACTTG GTTGTCAACCG CTGAATGTCA
 8941 CAGGGCGCGA TAGGTGCTAC GACCTAAGTT TAGGACTGCC GGCTGACGCC GGCAGGTTCG
 9001 ACTGGTCTT TGTGAACATT CACACGGAAAT TCAGAATCCA CCACTACCAAG CAGTGTGTCG
 9061 ACCACGCCAT GAAGCTGCCAG ATGCTTGGGG GAGATGCGCT ACCACTGCTA AAACCCGGCG
 9121 GCATCTTGAT GAGAGCTTAC GGATACGCCG ATAAAATCG CGAAGCCGTT GTTCCCTCCT
 9181 TAAGCAGAAA GTTCTCGTCT GCAAGAGTGT TGCGCCCGGA TTGTGTCAAG AGCAATACAG
 9241 AAGTGTCTT GCTGTTCTCC AACTTTGACA ACGGAAAGAG ACCCTCTACG CTACACCAAGA
 9301 TGAATACCAA GCTGAGTGCCT GTGTATGCCG GAGAACCCAT GCACACGGCC GGGTGTGCAC
 9361 CATCCTACAG AGTTAAGAGA GCAGACATAG CCACGTGCAC AGAAGCGGCT GTGGTTAACG-

FIGURE 29d

67/240

9421 CAGCTAACGC CGGTGGAAC GTAGGGGATG GCGTATGCAG GGCCGTGGCG AAGAAATGGC
 9481 CGTCAGCCCT TAAGGGAGCA GCAACACCGAG TGGGCACAAT TAAAACAGT ATGTGCGGCT
 9541 CGTACCCCGT CATCCACGCT GTAGCGCTA ATTTCCTCTGC CACGACTGAA GCGGAAGGG
 9601 ACCGCGAATT GGCGCTGTC TACCGGGCAG TGGCCGCCGA AGTAAACAGA CTGTCACTGA
 9661 GCAGCGTAGC CATCCCCTG CTGTCCACAG GAGTGGTCAG CGGGGAAAGA GATAGGCTGC
 9721 AGCAATCCCT CAACCATCTA TTACACAGCAA TGGACGCCAC GGACGCTGAC GTGACCATCT
 9781 ACTGCAGAGA CAAAAGTTGG GAGAAGAAAA TCCAGGAAGC CATTGACATG AGGACGGCTG
 9841 TGGAGTTGCT CAATGATGAC GTGGAGCTGA CCACAGACTT GGTGAGAGTG CACCCGGACA
 9901 GCAGCCTGGT GGGTCGTAAG GGCTACAGTA CCACGTACGG GTGCGCTGTAC TCGTACTTTG
 9961 AAGGTACGAA ATTCAACCGAG GCTGCTATTG ATATGGAGA GATACTGACG TTGTGGCCCA
 10021 GACTGCAAGA GGCAAACGAA CAGATATGCC TATAGCGCT GGGCGAAAACA ATGGACAAACA
 10081 TCAGATCCAA ATGTCGGGT AACGATTCCG ATTCACTAAC ACCTCCAGG ACAGTGCCT
 10141 GCCTGTGCCG CTACGCAATG ACAGCAGAAC GGATGCCCG CCTTAGGTCA CACCAAGTTA
 10201 AAAGCATGGT GGTTTGTCA TCTTTTCCCC TCCCGAAATA CCATGTAGAT GGGGTGCAGA
 10261 AGGTAAGTG CGAGAAGGTT CTCCCTGTTG ACCCGACGGT ACCTTCAGTG GTTACTCCGC
 10321 GGAAGTATGC CGCATCTACG ACGGACCAC TCCACTGCCA GCGATACCAT GTGCGTACCC AGTTTGCAGT
 10381 ACTGGACCAC CGACTCGTCT TCCACTGCCA GCGATACCAT GTGCGTACCC AGTTTGCAGT
 10441 CGTGTGACAT CGACTCGATC TACGAGCCAA TGGCTCCCAT AGTAGTGACG GCTGACGTAC
 10501 ACCCTGAACC CGCAGGCATC GCGGACCTGG CGGCAGATGT GCACCCCTGAA CCCGCAGACC
 10561 ATGTGGACCT GGAGAACCCG ATTCCCTCAC CGCGCCCGAA GAGAGCTGCA TACCTTGCCCT
 10621 CCCGCGCCGC GGAGCGACCG GTGCCGGCGC CGAGAAAGCC GACGCCCTGCC CCAAGGACTG
 10681 CGTTTAGGAA CAAGCTGCCCT TTGACGTTG CGACTTTGAG CGAGCACGAG GTCGATGCCGT
 10741 TGGCCTCCGG GATTACTTC GGAGACTTCG ACGACGCTCCT GCGACTAGGC CGCGCGGGTG
 10801 CATATATTTT CTCCCTGGAC ACTGGCAGCG GACATTTACA ACAAAAATCC GTTAGGCAGC
 10861 ACAATCTCCA GTGCGCACAA CTGGATGCGG TCCAGGAGGA GAAAATGTAC CCGCCAAAAT
 10921 TGGATACTGA GAGGGAGAAG CTGTTGCTGC TGAAAATGCA GATGCAACCCA TCGGAGGCTA
 10981 ATAAGAGTCG ATACCACTCT CGAAAGTGG AGAACATGAA AGCCACGGTG GTGGACAGGC
 11041 TCACATCGGG GGCCAGATTG TACACGGGAG CGGACGTAGG CGGCATACCA ACATACGCCG
 11101 TTCGGTACCC CGGCCCCGTG TACTCCCTA CGGTGATCGA AAGATTCTCA AGCCCCGATG
 11161 TAGCAATCGC AGCGTCAAC GAATACCTAT CCAGAAATTAA CCCAACAGTG GCGTCGTAC
 11221 AGATAACAGA TGAATACGAC GCATACTTGG ACATGGTTGA CGGGTCGGAT AGTTGCTTGG
 11281 ACAGAGCGAC ATTCTGCCCG GCGAAGCTCC GGTGCTACCC GAAACATCAT GCGTACCAACC
 11341 AGCCGACTGT ACGGCAGTGC GTCCCGTCAAC CCTTTAGAA CACACTACAG AACGTGCTAG
 11401 CGGCTGCCAC CAAGAGAAC TGCAACGTCA CGCAAATGCG AGAAACTACCC ACCATGGACT
 11461 CGGCAGTGT CAACGTGGAG TGCTTCAAGC GCTATGCCGCTCCTGGAGAA TATTGGGAAG
 11521 AATATGCTAA ACAACCTATC CGGATAACCA CTGAGAACAT CACTACCTAT GTGACCAAAT
 11581 TGAAAGGCC GAAAGCTGCT GCCTTGTTCG CTAAGACCCA CAACTTGGTT CGCTGCGAGG
 11641 AGGTTCCCAT GGACAGATTG ACGGTGACA TGAAACGAGA TGTCAAAGTC ACTCCAGGG
 11701 CGAAACACAC AGAGGAAAGA CCCAAAGTCC AGGTAATTCA AGCAGCGGAG CCATTGGCGA
 11761 CCGCTTACCT GTGCGGCATC CACAGGGAAAT TAGTAAGGAG ACTAAATGCT GTGTTACGCC
 11821 CTAACGTGCA CACATTGTTT GATATGTGG CGAAGACATT TGACGCGATC ATCGCCTCTC
 11881 ACTTCCACCC AGGAGACCCG GTTCTAGAGA CGGACATTGC ATCATTGAC AAAAGCCAGG
 11941 ACGACTCCTT GGCTTCAACA GGTTTAATGA TCCTCGAAGA TCTAGGGGTG GATCAGTAC
 12001 TGCTGGACTT GATCGAGGCA GCCTTGGGG AAATATCCAG CTGTCACCTA CCAACTGGCA
 12061 CGCGCTTCAA GTTCCGGAGCT ATGATGAAAT CGGGCATGTT TCTGACTTTG TTTATTAACA
 12121 CTGTTTGAA CATCACCATA GCAAGCAGGG TACTGGAGCA GAGACTCACT GACTCCGCCT
 12181 GTGCGGCCCTT CATCGGGAC GACAACATCG TTCACGGAGT GATCTCCGAC AAGCTGATGG
 12241 CGGAGAGGTG CGCGTGTGG GTCAACATGG AGGTGAAGAT CATTGACGCT GTCATGGCG
 12301 AAAAACCCCC ATATTTGT GGGGGATTCA TAGTTTTGA CAGCGTCACA CAGACCCCT
 12361 GCCGTGTTTC AGACCCACTT AAGGCCCTGT TCAAGTTGGG TAAGCCGCTA ACAGCTGAAG
 12421 ACAAGCAGGA CGAAGACAGG CGACGAGCAC TGAGTGACGA GGTT

FIGURE 29E

68/240

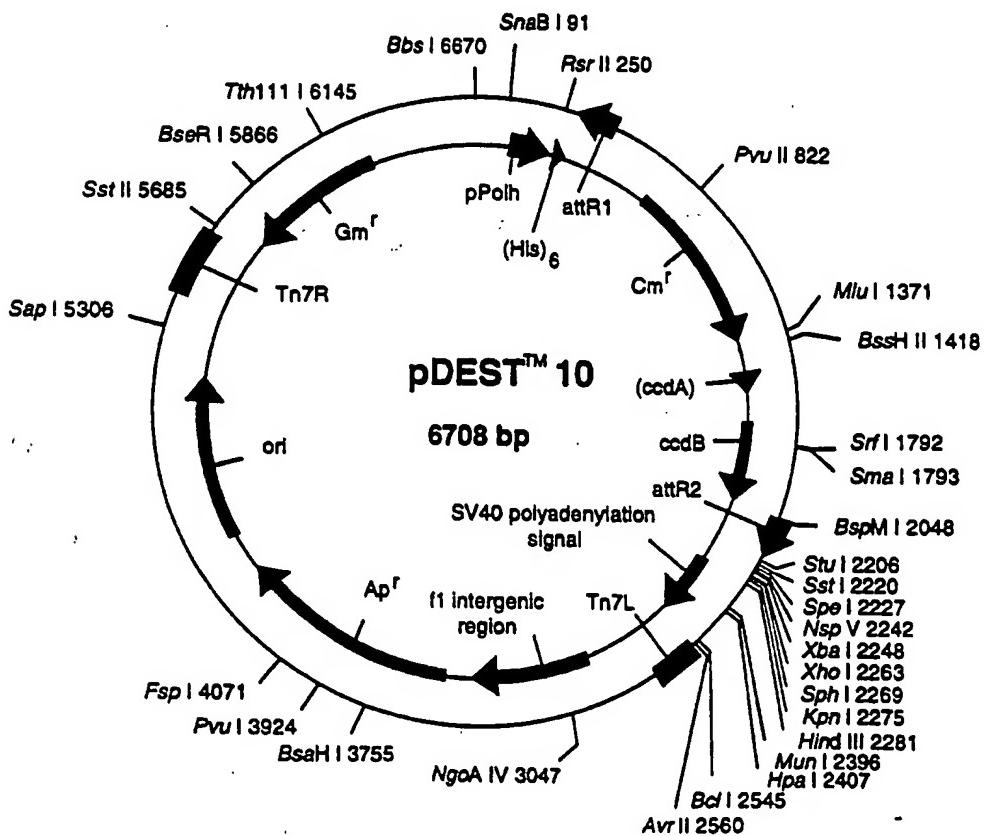
Figure 30A: pDEST10 Polyhedron Promoter with N-His6,
Baculovirus Transfer Plasmid

154 → mRNP from polyhedrin promoter
 aaa taa gta ttt tac tgc ttt cgt aac agt ttt gta ata aaa aaa cct ata
 ttt att cat aaa atg aca aaa gca ttg tca aaa cat tat ttt ttt gga tat

205 eat att ceg gat tat tea tac cgt ccc acc atc ggg egc gga tct egg tcc
 tta taa ggc cta ata agt atg gca ggg tgg tag ccc gcg cct aga gcc agg

256 Met Ser Tyr Tyr His His His His Asp Tyr Arg Ile Pro
 gaa acc atg tgg tac tac cat cac cat cac cat cac gat tac gat atc cca
 ctt tgg tac agc atg atg gta gtg gta gtg gta gtg cta atg cta tag ggt

307 Thr Thr Glu Asn Leu Tyr Phe Gln^{Gly} Ile Thr Ser Leu Tyr Lys Lys
 acg acc gaa aac ctg tat ttt cag ggc atc ~~aca~~ ~~act~~ ~~tgg~~ ~~tcc~~ ~~atc~~ ~~aaa~~ ~~gtc~~ ~~ccg~~ ~~tag~~ ~~tgt~~ ~~tca~~ ~~aac~~ ~~atg~~ ~~ttt~~ ~~tcc~~ ~~ogg~~
attR1 Int



69/240

pDEST10 6708 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
23..152	Ppolh
461..337	attR1
711..1370	CmR
1490..1574	inactivated ccdA
1712..2017	ccdB
2058..2182	attR2
3394..4369	ampR
4510..5164	ori
5658..62	genR

1 CCCCCGATGA AGTGGTTCGC ATCCCTGGT TTCTGGAAGG CGAGGCATCGT TTGTTGCC
 61 AGGACTCTAG CTATAGTTCT AGTGGTTGGC TACGTATACT CCGGAATATT AATAGATCAT
 121 GGAGATAATT AAAATGATAA CCATCTCGCA AATAAAATAAG TATTTTACTG TTTTCGTAAC
 181 AGTTTTGTAA TAAAAAAAACC TATAAAATATT CCGGATTATT CATACCGTCC CACCATCGGG
 241 CGCGGATCTC GGTCCGAAAC CATGTCGTAC TACCATCACC ATCACCATCA CGATTACGAT
 301 ATCCCAACGA CCGAAAACCT GTATTTTCAG GGCATCACAA GTTTGTACAA AAAAGCTGAA
 361 CGAGAACGT AAAATGATAT AAATATCAAT ATATTAATT AGATTTTGCA TAAAAAACAG
 421 ACTACATAAT ACTGTAAAAC ACAACATATC CAGTCACTAT GGCGGCCGCT AAGTTGGCAG
 481 CATCACCCGA CGCACTTTGC GCCGAATAAA TACCTGTGAC GGAAGATCAC TTCGCAGAAT
 541 AAATAAATCC TGTTGTCCTC GTTGTATACCG GGAAGCCCTG GGCCAACCTTT TGGCGAAAAT
 601 GAGACGTTGA TCGGCACGTA AGAGGTTCCA ACTTTCACCA TAATGAAATA AGATCACTAC
 661 CGGGCGTATT TTTTGAGTTA TCGAGATTTC CAGGAGCTAA GGAAGCTAAA ATGGAGAAAA
 721 AAATCACTGG ATATACCACC GTTGTATAT CCCAATGGCA TCGTAAAGAA CATTITGAGG
 781 CATTTCAGTC AGTTGCTCAA TGTACCTATA ACCAGACCGT TCAGCTGGAT ATTACGGCCT
 841 TTTTAAAGAC CGTAAAGAAA AATAAGCACA AGTTTTATCC GGCCCTTATT CACATTCTT
 901 CCCGCCTGAT GAATGCTCAT CCGGAATTCC GTATGGCAAT GAAAGACGGT GAGCTGGTGA
 961 TATGGGATAG TGTTCACCCCT TGTTACACCG TTTTCCATGA GCAAAACTGAA ACGTTTCAT
 1021 CGCTCTGGAG TGAATACCAAC GACGATTTC GGCAGTTCT ACACATATAT TCGCAAGATG
 1081 TGGCGTGTAA CGGTGAAAAC CTGGCCTATT TCCCTAAAGG GTTTATTGAG AATATGTTT
 1141 TCGTCTCAGC CAATCCCTGG GTGAGTTCA CCAGTTTGA TTTAAACGTG GCCAATATGG
 1201 ACAACTTCTT CGCCCCCGTT TTCACCATGG GCAAAATATTA TACGCAAGGC GACAAGGTGC
 1261 TGATGCCGCT GGGGATTTCAG GTTCATCATG CCGCTGTGA TGGCTTCCAT GTCGGCAGAA
 1321 TGCTTAATGA ATTACAACAG TACTGCGATG AGTGGCAGGG CGGGGGCGTAA ACGCGTGGAT
 1381 CGGGCTTACT AAAAGCCAGA TAACAGTATC CGTATTGCG CGCTGATTTC TGCGGTATAA
 1441 GAATATATAC TGATATGTAT ACCCGAAAGTA TGTCAAAAG AGGTGTGCTA TGAAGCAGCG
 1501 TATTACAGTG ACAGTTGACA GCGACAGCTA TCAGTTGCTC AAGGCATATA TGATGTCAT
 1561 ATCTCCGGTC TGTTAAGCAC AACCATGCAG AATGAAGCCC GTCGTCTGCC TGCCGAACGC
 1621 TGGAAAGCGG AAAATCAGGA AGGGATGGCT GAGGTGCCCG GTTTTATTGA AATGAACCGC
 1681 TCTTTTGCTG ACCAGAACAG GGACTGGTGA AATGCAGTTT AAGGTTTACA CCTATAAAAG
 1741 AGAGAGCCGT TATCGTCTGT TTGTGGATGT ACAGAGTGAT ATTATTGACA CGCCCGGGCG
 1801 ACGGATGGTC ATCCCCCTGG CCAGTCACG TCTGCTGTCA GATAAAAGTCT CCCGTGAACCT
 1861 TTACCCGGTG GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCACCG ATATGGCCAG
 1921 TGTGCCGGTC TCCGTTATCG GGGAAAGAAGT GGCTGATCTC AGCCACCGCG AAAATGACAT
 1981 CAAAAACGCC ATTAACCTGA TGTTCGGGG AATATAAAATG TCAGGCTCCC TTATACACAG
 2041 CCAGTCGCA GGTGACCAT AGTGACTGGA TATGTTGTGT TTACAGTAT TATGTAGTCT
 2101 GTTTTTATG CAAAATCTAA TTAAATATAT TGATATTAT ATCATTTTAC GTTTCTCGTT
 2161 CAGCTTCTT GTACAAAGTG GTGATGCCAT GGATCCGAA TTCAAAGGCC TACGTGACG
 2221 AGCTCAACTA GTGCCGGCCG TTTCGAATCT AGAGCCTGCA GTCTCGAGGC ATGCGGTACC
 2281 AAGCTTGTGAGAAGTACTA GAGGATCATA ATCAGCCATA CCACATTGT AGAGGTTTTA
 2341 CTTGCTTAA AAAACCTCCC ACACCTCCCC CTGAACCTGA AACATAAAAT GAATGCAATT
 2401 GTTGTGTTA ACTTGTTAT TGCAGTTAT AATGGTTACA AATAAAGCAA TAGCATCACA
 2461 AATTCACAA ATAAAGCATT TTTTCACTG CATTCTAGTT GTGGTTGTC CAAACTCATC
 2521 AATGTATCTT ATCATGTCTG GATCTGATCA CTGCTTGAGC CTAGGAGATC CGAACCGAGAT
 2581 AAGTGAATC TAGTTCCAAA CTATTTGTC ATTTTTAATT TTCGTATTAG CTTACGACGC-

Figure 30B

70/240

2641 TACACCCAGT TCCCATCTAT TTTGTCACTC TTCCCTAAAT AATCCTAAA AACTCCATT
 2701 CCACCCCTCC CAGTTCCCAA CTATTTGTC CGCCCACAGC GGGGCATTTT TCTTCCTGTT
 2761 ATGTTTTAA TCAAACATCC TGCCAATCC ATGTCACAAA CGTCATCTT CGGCTACTTT
 2821 TTCTCTGTCA CAGAATGAAA ATTTTCTGT CATCTCTCG TTATTAATGT TTGTAATTGA
 2881 CTGAATATCA ACCCTTATTT GCAGCTGAA TGGCGATGG GACCGGCCCT GTAGCGCGC
 2941 ATTAAGCGCG GCGGGTGIGG TGGTTACCGC CAGCGTGACC GCTACACTTG CCAGCGCCCT
 3001 AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC CTTCCTCGCC ACGTTCGCCCG GCTTTCCCCG
 3061 TCAAGCTCTA AATCGGGGGC TCCCTTAGG GTTCCGATTT AGTGCCTTAC GGCACCTCGA
 3121 CCCCAAAAAA CTTGATTAGG GTGATGGTC ACGTAGTGGG CCATCGCCCT GATAGACGGT
 3181 TTTTCGCCCT TTGACGTTGG AGTCCACGTT CTITAATAGT GGACTCTTGT TCCAAAATCTGG
 3241 AACAAACACTC AACCCATCT CCGTCTATTG TTTTGATTTA TAAGGGATTT TGCCGATTTC
 3301 GGCCTATTGG TTAAAAAATG AGCTGATTG ACAAAATTT AACCGGAATT TAAACAAAAT
 3361 ATTAACGTTT ACAATTTCAG GTGGCACTT TCGGGAAAT GTGCGCGGAA CCCCTATTG
 3421 TTATTTTC TAAATACATT CAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT
 3481 GCTTCATAA TATTGAAAAA GGAAGAGTAT GAGTATTCA CATTTCGCG TGCCCTTAT
 3541 TCCCTTTTGC GCGCATTTC GCTTCCCTGT TTTTGCCTAC CCAGAAACGC TGGTGAAACT
 3601 AAAAGATGCT GAAGATCAGT TGGGTCACG AGTGGTTAC ATCGAACCTGG ATCTCAACAG
 3661 CGGTAAGATC CTTGAGAGTT TTGACCGGAGA AGAACGTTT CCAATGATGA GCACTTTAA
 3721 AGTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTCG
 3781 CCGCATAAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT
 3841 TACGGATGGC ATGACAGTAA GAGAATTATG CAGTCTGCC ATAACCATGA GTGATAAACAC
 3901 TGCGGCCAAC TTACTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTGCA
 3961 CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGAA CCGGAGCTGA ATGAAGCCAT
 4021 ACCAAACGAC GAGCGTGACA CCACGATGCC TGAGCAATG GCAACAAACGT TCGCCAAAC
 4081 ATTAACTGGC GAACTACTTA CTCTAGCTT CCGGCAACAA TTAATAGACT GGATGGAGGC
 4141 GGATAAAGTT GCAGGACCCAC TTCTGGCTC GGGCCCTCCG GCTGGCTGGT TTATTGCTGA
 4201 TAAATCTGGA GCGGGTGAGC GTGGGCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG
 4261 TAAGCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG
 4321 AAATAGACAG ATCGCTGAGA TAGGTGCCCTC ACTGATTAAG CATTGGTAAC TGTCAAGACCA
 4381 AGTTTACTCA TATATACTTT AGATTGATTG AAAACTTCAT TTTTAATTAA AAAGGATCTA
 4441 GGTGAAGATC CTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCTTCCA
 4501 CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTCTGCG
 4561 CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGGCGGA
 4621 TCAAGAGCTA CCAACTCTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACCAA
 4681 TACTGTCCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAAAACTCTG TAGCACCGCC
 4741 TACATACCTC GCTCTGCTAA TCCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG
 4801 TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC
 4861 GGGGGTTCG TGACACACAGC CCAGCTTGA GCGAACGACC TACACCGAAC TGAGATAAC
 4921 ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAAGGCGG ACAGGTATCC
 4981 GGTAAAGCGGC AGGGTGGAA CAGGAGAGCG CACGAGGGAG CTCCAGGGGG GAAACGCCCTG
 5041 GTATCTTAT AGTCTGTGCG GTTTTCGCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG
 5101 CTCGTCAGGG GGGGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCTTTT TACGGTTCT
 5161 GGCCTTTGTC TGGCTTTTG CTCACATGTT CTTTCTGCG TTATCCCTG ATTCTGTTGA
 5221 TAACCGTATT ACCGCCTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG
 5281 CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCTGATG CGGTATTTTC TCCTTACGCA
 5341 TCTGTGCGGT ATTCACACC GCAGACCAGC CGCGTAACCT GCGAAAATCG GTTACGGTTG
 5401 AGTAATAAAT GGATGCCCTG CGTAAGCGGG TGTGGCGGA CAATAAAGTC TTAAACTGAA
 5461 CAAAATAGAT CTAAACTATG ACAATAAAAGT CTTAAACTAG ACAGAAATAGT TGTAAACTGAA
 5521 AATCAGTCCA GTTATGCTGT GAAAAAGCAT ACTGGACTTT TTGTTATGGCT AAAGCAAAC
 5581 CTTCATTTC TGAAGTGCCTA ATTGCCGTC GTATTAAGA GGGCGTGGC CAAGGGCATG
 5641 GTAAAGACTA TATTCGCGGC GTTGTGACAA TTTACCGAAC AACTCCGCGG CCGGGAAAGCC
 5701 GATCTCGGCT TGAACGAATT GTTGGTGGC GGTACTTGGG TCGATATCAA AGTGCATCAC
 5761 TTCTTCCCGT ATGCCCAACT TTGTATAGAG AGCCACTGCG GGATCGTCAC CGTAATCTGC
 5821 TTGCACGTAG ATCACATAAG CACCAAGCGC GTGGGCTCA TGCTTGAGGA GATTGATGAG
 5881 CGCGGTGGCA ATGCCCTGCC TCCGGTGCCTC GCGGGAGACT GCGAGATCAT AGATATAGAT
 5941 CTCACTACGC GGCTGCTAA ACCTGGGAG AACGTAAGCC GCGAGAGCGC CAACAACCGC
 6001 TTCTTGGTCG AAGGCAGCAA GCGCGATGAA TGTCTTACTA CGGAGCAAGT TCCCGAGGTA
 6061 ATCGGAGTCC GGCTGATGTT GGGAGTAGGT GGCTACGTCT CGGAACCTCAC GACCGAAAAG

FIGURE 30C

6121 ATCAAGAGCA GCCCGCATGG ATTTGACTTG GTCAGGGCCG AGCCTACATG TGCAGATGAT
6181 GCCCATACTT GAGGCCACCTA ACTTTGTTTT AGGGCACTG CCCTGCTGCG TAACATCGTT
6241 GCTGCTGCGT AACATCGTT CTGCTCCATA ACATCAAACA TCGACCCACG GCGTAACCGCG
6301 CTTGCTGCTT GGATGCCCGA GGCATAGACT GTACAAAAAA ACAGTCATAA CAAGCCATGA
6361 AAACCGCCAC TGCGCCGTTA CCACCGCTGC GTTCGGTCAA GGTTCTGGAC CAGTTGCGTG
6421 AGCGCATACG CTACTTGAT TACAGTTAAC GAACCGAACAA GGCTTATGTC AACTGGGTTG
6481 GTGCCTTCAT CCGTTTCCAC GGTGTGCGTC ACCCGGCAAC CTTGGGCAGC AGCGAAGTCG
6541 AGGCATTTCGT GTCTGGCTG GCGAACGAGC GCAAGGTTTC GGTCTCCACG CATCGTCAGG
6601 CATTGGCGGC TTGCTGTTT TTCTACGGCA AGGTGCTGTG CACGGATCTG CCCTGGCTTC
6661 AGGAGATCGG AAGACCTCGG CCGTCGCGGC GCTTGGCGGT GGTGCTGA

FIGURE 30D

72/240

Figure 31A: pDEST11

Tet-regulated eukaryotic expression

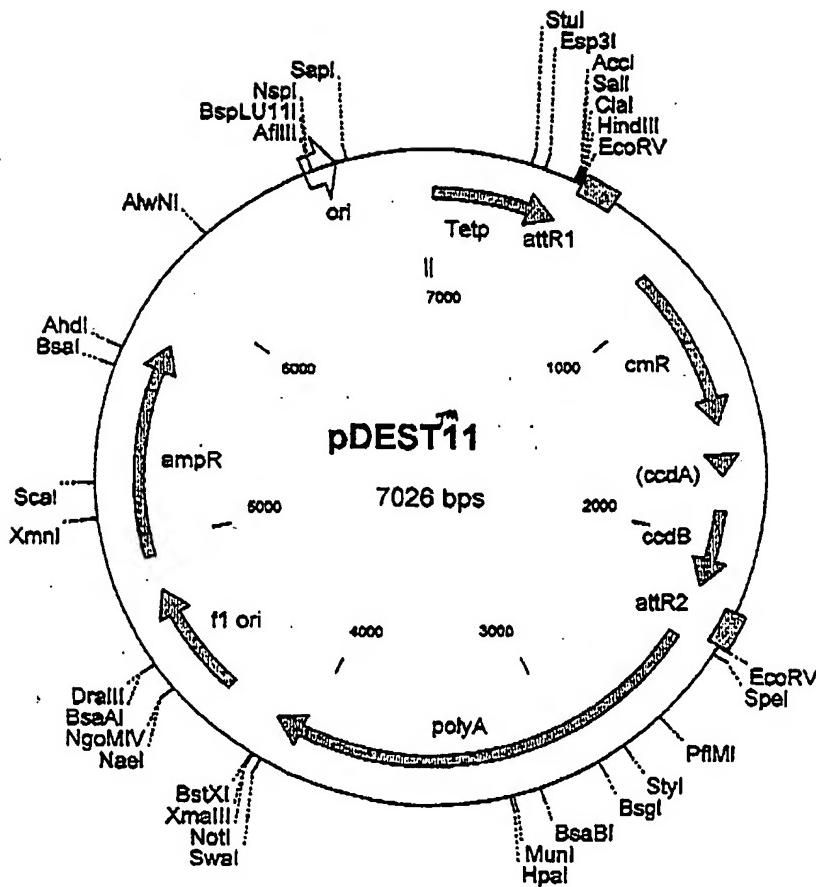
mRNA from CMV promoter (controlled by tetracycline)

358 tag tga acc gEc aga teg cct gga gac gcc aac cac get gtt tgg acc tcc
atc act tgg cag tct agc gga cct ctg cgg tag gtg cga caa aac tgg agg

409 ata gaa gac acc ggg acc gat cca gcc tcc ggg gcc ccc aat tcg agc tcg
tat ctt ctg tgg ccc tgg cta ggt cgg agg cgc cgg ggc tta agc tcg agc

460 gta ccc ggg gat cct cta gag tcg agg tcc acc gta tcc ata tcg ttc ada
cat ggg ccc cta gga gat ctc agc tcc agc tcc cat agc tat tcg acc tat

511 tca aca agt tcg aaa aac aaa gct gaa cga gaa acg taa dat gat abt aat
agt tgt tca aac atg ttt tct cca ctt act ccc tgc att tta cta dat tta



73/240

pDEST11 7026 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
4..479	Tet ^r ((Tet operator)7 and min hCMV promoter)
638..514	attR1
888..1547	CmR
1667..1751	inactivated ccdA
1889..2194	ccdB
2235..2359	attR2
2402..4132	polyA
4347..4803	f1 ori
4940..5797	ampR

1 CGAGTTTAC CACTCCCTATC AGTGATAGAG AAAAGTGAAG GTCGAGTTTA CCACTCCCTA
 61 TCAGTGATAG AGAAAAGTGA AAGTCGAGTT TACCACTCCC TATCAGTGAT AGAGAAAAGT
 121 GAAAGTCGAG TTTACCACTC CCTATCAGTG ATAGAGAAAA GTGAAAGTCG AGTTTACAC
 181 TCCCTATCAG TGATAGAGAA AAGTGAAGT CGAGTTTAC CACTCCCTATC AGTGATAGAG
 241 AAAAGTGAAG GTCGAGTTTA CCACTCCCTA TCAGTGATAG AGAAAAGTGA AAGTCGAGCT
 301 CGGTACCCGG GTCGAGTAGG CGTGTACGGT GGGAGGCCCTA TATAAGCAGA GCTCGTTAG
 361 TGAACCGTCA GATCGCCTGG AGACGCCATC CACGCTGTT TGACCTCCAT AGAACGACACC
 421 GGGACCGATC CAGCCTCCGC GGCCCCGAAT TCGAGCTCGG TACCCGGGGA TCCTCTAGAG
 481 TCGAGGTCGA CGGTATCGAT AAGCTTGATA TCAACAAGTT TGTACAAAAA AGCTGAACGA
 541 GAAACGTAAA ATGATATAAA TATCAATATA TAAATTAGA TTGTCATAA AAAACAGACT
 601 ACATAATACT GTAAAACACA ACATATCCAG TCACTATGGC GGCGCTAAG TTGGCAGCAT
 661 CACCCGACGC ACTTTGCGCC GAATAATAC CTGTGACGGA AGATCACTTC GCAGAATAAA
 721 TAAATCCTGG TGTCCTGTT GATACCGGGA AGCCCTGGGC CAACTTTGG CGAAAATGAG
 781 ACGTTGATCG GCACGTAAGA GGTTCCAATC TTCACCATAA TGAAATAAGA TCACTACCGG
 841 GCGTATTTTG TGAGTTATCG AGATTTTCAG GAGCTAAGGA AGCTAAAATG GAGAAAAAAA
 901 TCACTGGATA TACCAACCGTT GATATATCCC AATGGCATCG TAAAGAACAT TTTGAGGCAT
 961 TTCACTGAGT TGCTCAATGT ACCTATAACC AGACCGTTCA GCTGGATATT ACGGCCCTTTT
 1021 TAAAGACCGT AAAGAAAAAT AAGCACAAGT TTATCCGGC CTTTATTTCAC ATTCTTGCCC
 1081 GCCTGATGAA TGCTCATCCG GAATTCCGTA TGGCAATGAA AGACGGTGAG CTGGTGTAT
 1141 GGGATAGTGT TCAACCTTGT TACACCGTT TCCATGAGCA AACTGAAACG TTTTCATCGC
 1201 TCTGGAGTGA ATACCACGAC GATTTCCGGC AGTTTCTACA CATATATTTCG CAAGATGTGG
 1261 CGTGTACGG TGAAAACCTG GCCTATTTC CCAAAGGGTT TATTGAGAAT ATGTTTTTCG
 1321 TCTCAGCCAA TCCCTGGGTG AGTTTCACCA GTTTGATTT AAACGTGGCC AATATGGACA
 1381 ACTTCTTCGC CCCCGTTTTC ACCATGGGCA AATATTATAC GCAAGGCGAC AAGGTGCTGA
 1441 TGCCGCTGGC GATTCAAGGTT CATCATGCCG TCTGTGATGG CTTCCATGTC GGCAGAATGC
 1501 TTAATGAATT ACAACAGTAC TGCGATGAGT GGCAGGGCGG GGCAGTAAAGA TCTGGATCCG
 1561 GCTTACTAAA AGCCAGATAA CAGTATGCGT ATTTGCGCG TGATTTTGCG GGTATAAGAA
 1621 TATATACTGTA TATGTATACC CGAAGTATGT CAAAAAGAGG TGTGCTATGA AGCAGCGTAT
 1681 TACAGTGACA GTTGACAGCG ACAGCTATCA GTTGCTCAAG GCATATATGA TGTCAATATC
 1741 TCCGGTCTGG TAAGCACAAAC CATGCAGAAT GAAGCCCGTC GTCTGCGTGC CGAACGCTGG
 1801 AAAGCGGAAA ATCAGGAAGG GATGGCTGAG GTGGCCCGGT TTATTGAAAT GAACGGCTCT
 1861 TTGCTGACG AGAACAGGGG CTGGTGAAT GCAGTTAAG GTTACACCT ATAAAAGAGA
 1921 GAGCCGTTAT CGTCTGTTTG TGGATGTACA GAGTGTGATATT ATTGACACGC CCGGGCGACG
 1981 GATGGTGATC CCCCTGGCCA GTGCACGCTCT GCTGTCAGAT AAAGTCTCCC GTGAACCTTTA
 2041 CCCGGTGGTG CATATCGGGG ATGAAAGCTG GCGCATGATG ACCACCGATA TGGCCAGTGT
 2101 GCCGGTCTCC GTTATCGGGG AAGAAGTGGC TGATCTCAGC CACCGCGAAA ATGACATCAA
 2161 AAACGCCATT AACCTGATGT TCTGGGAAT ATAAATGTCA GGCTCCCTTA TACACAGCCA
 2221 GTCTGCGAGT CGACCATAGT GACTGGATAT GTTGTGTTTT ACAGTATTAT GTAGTCTGTT
 2281 TTTTATGCAA AATCTAATTT AATATATTGA TATTTATATC ATTTTACGTT TCTCGTTCA
 2341 CTTTCTTGTG CAAAGTGGTT GATATCGAAAT TCCTGCAGCC CGGGGGATCC ACTAGTTCTA
 2401 GAGCACTGCG ATGAGTGGCA GGGCGGGCG TAATTTTTT AAGGCAGTTA TTGGTGCCCT
 2461 TAAACGCCCTG GTGCTACGCC TGAATAAGTG ATAATAAGCG GATGAATGCG AGAAAATTGCG
 2521 CGGATCTTGT TGAAGGAACC TTACTTCTGT GGTGTGACAT AATTGGACAA ACTACCTACA-

FIGURE 31B

76/240

2581 GAGATTAAAGCTCTAAGGT AAAATATAAAA TTTTTAAGTG TATAATGTGT TAAACTACTG
 2641 ATTCTAATTG TTGTGTATT TTAGATTCCA ACCTATGGAA CTGATGAATG GGAGCAGTGG
 2701 TGGAAATGCCT TTAATGAGGA AAACCTGTT TGCTCAGAAG AAATGCCATC TAGTGATGAT
 2761 GAGGCTACTG CTGACTCTCA ACATTCTACT CCTCCAAAAA AGAAGAGAAA GGTAGAAC
 2821 CCCAAGGACT TTCCCTTCAGA ATTGCTAAGT TTTTGAGTC ATGCTGTGTT TAGTAATAGA
 2881 ACTCTTGCTT GCTTTGCTAT TTACACCACA AAGGAAAAAG CTGACTGCT ATACAAGAAA
 2941 ATTATGGAAA AATATTCTGT AACCTTATA AGTAGGCATA ACAGTTATAA TCATAACATA
 3001 CTGTTTTTC TTACTCCACA CAGGCATAGA GTGCTGCTA TTAATAACTA TGCTCAAAA
 3061 TTGTGTACCT TTAGCTTTT AATTGTTAAA GGGGTTAATA AGGAATATTG GATGTATAGT
 3121 GCCTTGACTA GAGATCATAA TCAGCCATAC CACATTGTA GAGGTTTAC TTGCTTAAA
 3181 AAACCTCCCCA CACCTCCCCC TGAACCTGAA ACATAAAATG AATGCAATTG TTGTTGTTAA
 3241 CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGCAAT AGCATCACAA ATTCACAAA
 3301 TAAAGCATTG TTTCACTGC ATTCTAGTT TGTTGTTGAA AACCTCATCA ATGTATCTTA
 3361 TCATGCTCGG ATCCCCAGGA AGCTCCTCTG TGCTCTCATA AACCCCTAAC CCCTCTACTT
 3421 GAGAGGACAT TCCAATCATA GGCTGCCAT CCACCCCTG TGCTCTCCTG TTAATTAGGT
 3481 CACTTAACAA AAAGGAAATT GGGTAGGGGT TTTCACAGA CCGCTTTCTA AGGGTAATT
 3541 TAAATATCT GGGAAAGTCCC TTCCACTGCT GTGTTCCAGA AGTGTGGTA AACAGCCCAC
 3601 AAATGTCAAC AGCAGAAACA TACAAGCTGT CAGCTTGCA CAAGGGCCCA ACACCCCTGCT
 3661 CATCAAGAAG CACTGTGGTT GCTGTGTTAG TAATGTGAA AACAGGAGGC ACATTTCCC
 3721 CACCTGTGTA GGGTTCAAAA TATCTAGTTG TTTCATTTT ACTTGGATCA GGAACCCAGC
 3781 ACTCCACTGG ATAAGCATTAA TCCTTATCCA AAACAGCCTT GTGGTCAGTG TTCATCTGCT
 3841 GACTGTCAAC TGTTAGCATTT TTGGGGTTA CAGTTTGAGC AGGATATTG GTCTGTAGT
 3901 TTGCTAACAC ACCCTGCAGC TCCAAAGGTT CCCCACCAAC AGCAAAAAAA TGAAAATTG
 3961 ACCCTTGAAT GGGTTTCCCA GCACCATTTT CATGAGTTT TTGTTGCTCCCT GAATGCAAGT
 4021 TTAACATAGC AGTTACCCCA ATAACCTCAG TTTTAACTAGT AACAGCTTCC CACATCAAAA
 4081 TATTTCCACA GGTTAAAGTCC TCATTTAAAT TAGGCAAAGG AATTGCTCTA GAGCGGCCGC
 4141 CACCGGGGTG GAGCTCCAAT TCGCCCTATA GTGAGTCGTA TTACGCGCGC TCACTGGCCG
 4201 TCGTTTACA ACGTCGTGAC TGGGAAAACC CTGGCGTTAC CCAACTTAAAT CGCCTTGCAG
 4261 CACATCCCCC TTTCGCCAGC TGGCGTAATA GCGAAGAGGC CGCACCAGAT CGCCCTTCCC
 4321 AACAGTTGCG CAGCCTGAAT GGCGAATGGG ACGCGCCCTG TAGCGCGCA TTAAGCGCG
 4381 CGGGGTGTTGGT GGTTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA GCGCCCGCTC
 4441 CTTTCGCTTT CTTCCCTTCC TTTCCTGCCA CGTCGCCGG CTTTCCCCGT CAAGCTCTAA
 4501 ATCGGGGGCT CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC CCCAAAAAAC
 4561 TTGATTAGGG TGATGGTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT TTTCGCCCTT
 4621 TGACGTTGGA GTCCACGTT TTTAATAGTG GACTCTGTT CCAAACCTGGA ACAACACTCA
 4681 ACCCTATCTC GGTCTATTCT TTGATTAT AAGGGATTG GCGATTTCG GCCTATTGGT
 4741 TAAAAAATGAA GCTGATTTAA CAAAAATTAA ACGCGAATTAA TAACAAAATA TTAACGCTTA
 4801 CAATTAGGT GGCACTTTTC GGGGAATGTC GCGCGGAACC CCTATTGTT TATTTTCTA
 4861 AATACATTCA AATATGTATC CGCTCATGAG ACAATAACCC TGATAAAATGC TTCAATAATA
 4921 TTGAAAAAGG AAGAGTATGA GTATTCAACA TTTCCTGTC GCCCTTATTG CTTTTTTGTC
 4981 GGCATTTGCT CTTCTGTT TTGCTCACCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA
 5041 AGATCAGTTG GGTGACGAG TGGGTTACAT CGAAGCTGGAT CTCAACAGCG GTAAGATCCT
 5101 TGAGAGTTT CGCCCCGAAG AACGTTTCC AATGATGAGC ACTTTAAAG TTCTGCTATG
 5161 TGGCGCGTA TTATCCCGTA TTGACGCCGG GCAAGAGCAA CTCGGTCGCC GCATACACTA
 5221 TTCTCAGAAT GACTGGTTG AGTACTCACC AGTCACAGAA AAGCATCTTA CGGATGGCAT
 5281 GACAGTAAGA GAATTATGCA GTGCTGCCAT AACCATGAGT GATAACACTG CGGCCAACTT
 5341 ACTCTGACA ACGATCGGAG GACCGAAGGA GCTAACCGCT TTTTGACACA ACATGGGGA
 5401 TCATGTAACCT CGCCCTTGATC GTTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGACGA
 5461 GCGTGACACC ACGATGCCTG TAGCAATGGC AACAAACGTT CGCAAACATAT TAACTGGCGA
 5521 ACTACTTACT CTAGCTTCCC GGCAACAATT AATAGACTGG ATGGAGGCGG ATAAAGTTGC
 5581 AGGACCACTT CTGCGCTCGG CCCTTCCGGC TGGCTGGTTT ATTGCTGATA AATCTGGAGC
 5641 CGGTGAGCGT GGGTCTCGCG GTATCATTGC AGCACTGGGG CCAGATGGTA AGCCCTCCCG
 5701 TATCGTAGTT ATCTACACGA CGGGGAGTCA GGCAACTATG GATGAACGAA ATAGACAGAT
 5761 CGCTGAGATA GGTGCTCAC TGATTAAGCA TTGTTAAGTG TCAGACCAAG TTACTCTATA
 5821 TATACTTTAG ATTGATTAA AACTTCATT TTAATTAAA AGGATCTAGG TGAAGATCCT
 5881 TTTTGATAAT CTCATGACCA AAATCCCTTA ACGTGAGTTT TCGTTCCACT GAGCGTCAGA
 5941 CCCCGTAGAA AAGATCAAAG GATCTCTTGT AGATCCTTTT TTCTGCGCG TAATCTGCTG
 6001 CTTGCAAACA AAAAAACCAC CGCTACCAAGC GGTGGTTGT TTGCGGGATC AAGAGCTACC-

FIGURE 3/C

75/240

6061 AACTTTTTT CCGAAGGTAA CTGGCTTCAG CAGAGCGCAG ATACCAAATA CTGTCCTTCT
6121 AGTGTAGCG TAGTTAGGCC ACCACTTCAA GAACTCTGTA GCACCGCCTA CATAACCTCGC
6181 TCTGCTAATC CTGTTACCAAG TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC TTACCGGGTT
6241 GGACTCAAGA CGATAGTTAC CGGATAAGGC GCAGCGGTG GGCTGAACGG GGGGTTCGTG
6301 CACACAGCCC AGCTTGGAGC GAACGACCTA CACCGAACTG AGATAACCTAC AGCGTGAGCT
6361 ATGAGAAAAGC GCCACGCTTC CCGAAGGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCAG
6421 GGTCGGAACA GGAGAGCGCA CGAGGGAGCT TCCAGGGGA AACGCCTGGT ATCTTTATAG
6481 TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT CGTCAGGGGG
6541 GCGGAGCCTA TGGAAAAACG CCAGCAACGC GGCCTTTTA CGGTTCCCTGG CCTTTTGCTG
6601 GCCTTTGCT CACATGTTCT TTCCCTGCGTT ATCCCCCTGAT TCTGTGGATA ACCGTATTAC
6661 CGCCTTGAG TGAGCTGATA CCGCTCGCCG CAGCCGAACG ACCGAGCGCA GCGAGTCAGT
6721 GAGCGAGGAA GCGGAAGAGC GCCCCATACG CAAACCGCCT CTCCCCGCGC GTTGGCCGAT
6781 TCATTAAATGC AGCTGGCACG ACAGGTTTCC CGACTGGAAA GCGGGCAGTG AGCGCAACGC
6841 AATTAATGTG AGTTAGCTCA CTCATTAGGC ACCCCAGGCT TTACACTTTA TGCTTCCGGC
6901 TCGTATGTTG TGTGGAATTG TGAGCGGATA ACAATTTCAC ACAGGAAACAA GCTATGACCA
6961 TGATTACGCC AAGCGCGCAA TTAACCCCTCA CTAAAGGGAA CAAAAGCTGG GTACCGGGCC
7021 CCCCT

FIGURE 31D

Figure 32A: pDEST12.2 CMV Promoter for Eukaryotic Expression, SV40 Promoter/ori for G418 Resistance

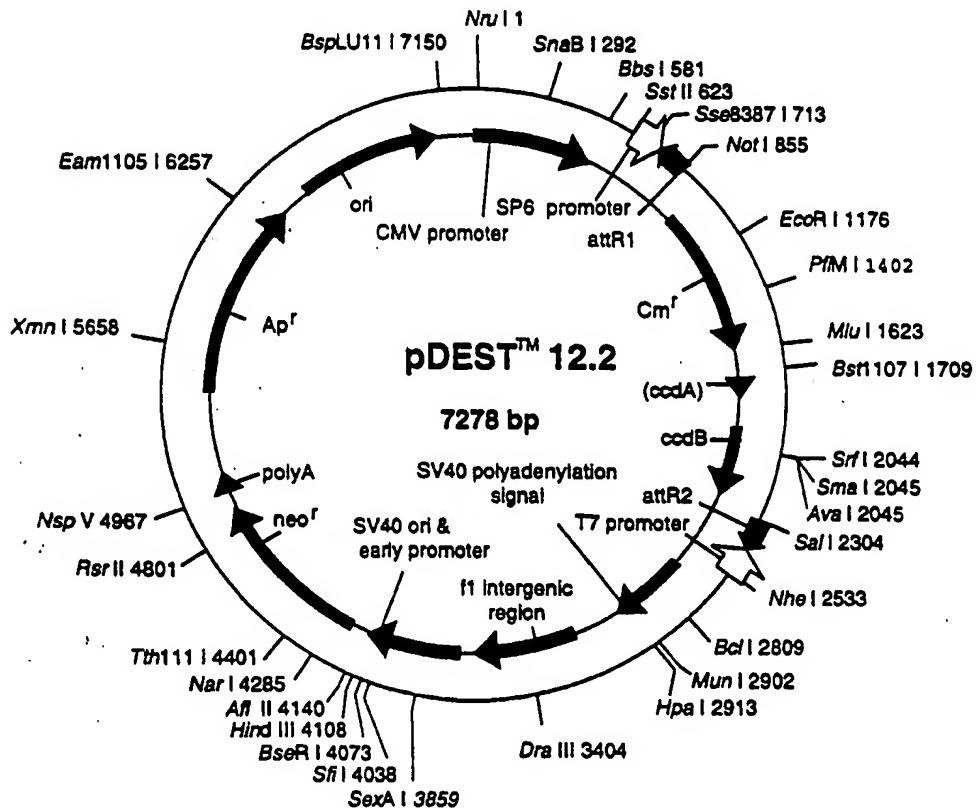
307 acc ~~gtc~~ aga tcc ~~cct~~ gga gac gcc atc cac ~~gtt~~ tgg acc tcc ata gaa
 tgg cag tct agc gga ~~cct~~ ctg cgg tag gtg cga caa aac tgg agg tat ctt

358 gac acc ggg acc gat cca gcc tcc gga ctc tag ~~cct~~ agg ccc cgg agg acg gga
 ctg tgg ccc tgg cta ggt cgg agg ~~cct~~ gag atc gga tcc ggc gec tgg cct

409 taa caa ttt cac aca gga aac agc tat gac cat tag gcc ttt gca aaa age
 att gtt aaa gtg tgt cct ttg tcg ata ctg gta atc cgg aaa cgt ttt tcg

450 tat tta ggt gac act ata gaa ggt acg cct gca ggt ~~acc~~ ggt ccc ~~aaa~~ ttc
 ata aat cca ctg tga tat ctt cca tgc gga cgt cca tgg cca ggc ctt aag

511 cca tca ~~aca~~ agt ~~tcc~~ tao ~~ada~~ ada ~~gtt~~ gaa cga gaa acg taa aat gat atc
~~gtt~~ tca aac atg ~~ttt~~ tbc cga ctc gct ctt tgc att ~~ata~~ cca ~~tat~~



77/240

pDEST12.2 7278 bp (rotated to position 3900)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
86..136	ori
220..742	CMV promoter
1059..935	attR1
1168..1827	CmR
1947..2031	inactivated ccdA
2169..2474	ccdB
2515..2639	attR2
2824..3186	small t & polyA
3310..3378	lac
4363..5157	neo
5680..6540	ampR

1 GGGGGCGGAA GCCTATGGAA AAACGCCAGC AACGCCGCCT TTTTACGGTT CCTGGCCTTT
 61 TGCTGGCCTT TTGCTCACAT GTTCTTCCT GCGTTATCCC CTGATTCTGT GGATAACCGT
 121 ATTACCGCCT TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG
 181 TCAGTGAGCG AGGAAGCGGA AGAGCTCGCG AATGCATGTC GTTACATAAC TTACGGTAAA
 241 TGGCCCGCCT GGCTGACCAG CCAACGACCC CCGCCCATTG ACGTCAATAA TGACGTATGT
 301 TCCCCATAGTA ACCCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA
 361 AACTGCCAAC TTGGCAGTAC ATCAAGTGT A CATATGCCA AGTACGCCCT CTATTGACGT
 421 CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCAGTAC ATGACCTTAT GGGACTTTCC
 481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGTGAC GGTGTTGGCA
 541 GTACATCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGAA TTTCCAAGTC TCCACCCCAT
 601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA AATGTGCTAA
 661 CAACTCCGCC CCATTGACGC AAATGGCCGG TAGGCGTGTAA CGGTGGGAGG TCTATATAAG
 721 CAGAGCTCGT TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT GTTTTGACCT
 781 CCATAGAAGA CACCGGGACC GATCCAGCCT CCGGACTCTA GCCTAGGCGG CGGGACGGAT
 841 AACAAATTCA CACAGGAAAC AGCTATGACC ATTAGGCCCT TGCAAAAGC TATTTAGGTG
 901 ACACTATAGA AGGTACGCCCT GCAGGTACCG GATCACAAGT TTGTACAAAA AAGCTGAACG
 961 AGAAACGTAA AATGATATAA ATATCAATAT ATTAAATTAG ATTTTGACATA AAAAACAGAC
 1021 TACATAATAC TGTAACACAC AACATATCCA GTCACATATGG CGGCCGCATT AGGCACCCCA
 1081 GGCTTTACAC TTTATGCTTC CGGCTCGTAT AATGTGTGGA TTTTGAGTTA GGATCCGTG
 1141 AGATTTTCAG GAGCTAAGGA AGCTAAAATG GAGAAAAAAA TCACTGGATA TACCACCGTT
 1201 GATATATCCC AATGGCATCG TAAAGAACAT TTTGAGGCAT TTCAGTCAGT TGCTCAATGT
 1261 ACCTATAACC AGACCGTTCA GCTGGATATT ACGGCCCTTT TAAAGACCGT AAAGAAAAAT
 1321 AAGCACAAGT TTATCCGGC CTTTATTCCAC ATTCTTGCCC GCCTGATGAA TGCTCATCCG
 1381 GAATTCGGTA TGGCAATGAA AGACGGTGAG CTGGTGATAT GGGATAGTGT TCACCCCTTGT
 1441 TACACCGTT TCCATGAGCA AACTGAAACG TTTTCATCGC TCTGGAGTGA ATACCACGAC
 1501 GATTTCCGGC AGTTTCTACA CATATATTCTG CAAGATGTGG CGTGTACGG TGAAAACCTG
 1561 GCCTATTTCCT CAAAGGGTT TATTGAGAAT ATGTTTTTCG TCTCAGCCAA TCCCTGGGTG
 1621 AGTTTACCCA GTTTGATTTT AAACGTGGCC AATATGGACA ACTTCTTCGC CCCCCTTTTC
 1681 ACCATGGGCA AATATTATAC GCAAGGGAC AAGGTGCTGA TGCCGCTGGC GATTCAAGGTT
 1741 CATCATGCCG TCTGTGATGG CTTCCATGTC GGCAGAATGC TTAATGAATT ACAACAGTAC
 1801 TGCAGTGAAGT GGCAGGGCGG GGCGTAAACG CGTGGATCCG GCTTACTAAA AGCCAGATAA
 1861 CAGTATGCGT ATTGCGCGC TGATTTTGC GGTATAAGAA TATATACTGA TATGTATACC
 1921 CGAAGTATGT CAAAAGAGG TGTGCTATGA AGCAGCGTAT TACAGTGACA GTTGACAGCG
 1981 ACAGCTATCA GTTGTCTCAAG GCATATATGA TGTCAATATC TCCGGTCTGG TAAGCACAAC
 2041 CATGCAGAAT GAAGCCCGTC GTCTGGTGC CGAACGCTGG AAAGGGAAA ATCAGGAAGG
 2101 GATGGCTGAG GTGCCCCGGT TTATTGAAAT GAACGGCTCT TTTGCTGACG AGAACAGGGA
 2161 CTGGTGAAT GCAGTTAAG GTTACACCT ATAAAAGAGA GAGCCGTTAT CGTCTGTTG
 2221 TGGATGTACA GAGTGATATT ATTGACACGC CGGGCGACG GATGGTGATC CCCCTGGCCA
 2281 GTGCACGTCT GCTGTCAAGT AAAGTCTCCC GTGAACTTA CCCGGTGGTG CATATCGGGG
 2341 ATGAAAGCTG GGCATGATG ACCACCGATA TGGCAGTGT GCGGGTCTCC GTTATCGGGG
 2401 AAGAAGTGGC TGATCTCAGC CACCGCGAAA ATGACATCAA AAACGCCATT AACCTGATGT-

FIGURE 32B

78/240

2461 TCTGGGAAAT ATAAATGTCA GGCTCCCTTA TACACAGCCA GTCTGCAGGT CGACCATACT
 2521 GACTGGATAT GTTGTGTTTT ACAGTATTAT GTAGTCTGTT TTTTATGCAA AATCTAATTT
 2581 AATATATTGA TATTTATATC ATTTCACGTT TCTCGTTCA CTTTCTTGTA CAAAGTGGTG
 2641 ATCGCGTGCA TGCGACGTCA TAGCTCTTC CCTATAGTGA GTCGTATTAT AAGCTAGGC
 2701 CTGGCCGTG 5' TTTTACAACG TCGTACTGG GAAAATGCT AGCTGGGAT CTTTGTGAAG
 2761 GAACCTTACT TCTGTGGGTG GACATAATTG GACAAACTAC CTACAGAGAT TTAAAGCTCT
 2821 AAGGTAATAA TAAAATTTT AAGTGTATAA TGTTGTTAAC TAGCTGCATA TGCTTGTGC
 2881 TTGAGAGTTT TGCTTACTGA GTATGATTAA TGAAAATATT ATACACAGGA GCTAGTGATT
 2941 CTAATTGTTT GTGTATTATA GATTACAGT CCCAAGGCTC ATTTCAAGGC CCTCAGTCCT
 3001 CACAGTCTGT TCATGATCAT AACAGCCAT ACCACATTG TAGAGGTTTT ACTTGCTTTA
 3061 AAAAACCTCC CACACCTCCC CCTGAACCTG AAACATAAAA TGAATGCAAT TGTTGTTGTT
 3121 AACTTGTAA TTGCACTTAA TAATGGTTAC AAATAAAGCA ATAGCATCAC AAATTCACA
 3181 AATAAACAT TTTTTCACT GCATTCTAGT TGTGGTTGT CCAAACACTCAT CAATGTATCT
 3241 TATCATGTCT GGATCGATCC TGCATTAATG AATCCGCCAA CGCGCGGGGA GAGGGGGTTT
 3301 GCGTATTGGC TGGCGTAATA GCGAAGAGGC CGCGACCGAT CGCCCTTCCC AACAGTTGCG
 3361 CAGCCTGAAT GGCAGATGGG ACGCGCCCTG TAGGGCGCA TTAAGCGCG CGGGTGTGGT
 3421 GGTTACGCG AGCGTGACCG CTACACTTGC CAGGCCCTA GCGCCCGCTC CTTTCGCTTT
 3481 CTTCCCTTCC TTTCTCGCCA CGTTCGCCGG CTTTCCCCGT CAAGCTCTAA ATCGGGGGCT
 3541 CCCTTTAGGG TTCCGATTTA GTGTTTACG GCACCTCGAC CCCAAAAAAC TTGATTAGGG
 3601 TGATGGTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT TTTCGCCCT TGACGTTGGA
 3661 GTCCACGTTC TTTAATAGTG GACTCTTGT CCAAACACTGA ACAACACTCA ACCCTATCTC
 3721 GGTCTATTCT TTTGATTATA AAGGGATTTT GCCGATTTCG GCCTATTGGT TAAAAAATGA
 3781 GCTGATTAA CAAATTTTA ACGCGAATT TAAACAAAATA TTAACGTTTA CAATTCGCC
 3841 TGATGCGTA TTTTCTCTT ACGCATCTGT GCGGTATTTC ACACCGCATA CGCGGATCTG
 3901 CGCAGCACCA TGGCCTGAAA TAAACCTCTGA AAGAGGAAC TGGTTAGGTA CCTTCTGAGG
 3961 CGGAAAGAAC CAGCTGTGGA ATGTGTGCA GTTAGGGTGT GGAAGAGTCCC CAGGCTCCCC
 4021 AGCAGGCAGA AGTATGCAA GCATGCATCT CAATTAGTCA GCAACCCAGGT GTGGAAAGTC
 4081 CCCAGGCTCC CCAGCAGGCA GAAAGTATGCA AAGCATGCA CTCAATTAGT CAGCAACCAT
 4141 AGTCCCCCCC CTAACTCCGC CCATCCCCCC CCTAACTCCG CCCAGTTCCG CCCATTCTCC
 4201 GCCCCATGGC TGACTAATT TTTTTATTAA TGCAGAGGCC GAGGCCGCGCT CGGCCTCTGA
 4261 GCTATTCCAG AAGTAGTGGAG GAGGCTTTT TGGAGGCCCTA GGCTTTGCA AAAAGCTTGA
 4321 TTCTTCTGAC ACAACAGTCT CGAACCTTAAG GCTAGAGCCA CCATGATTGA ACAAGATGGA
 4381 TTGCACGCGAG GTTCTCCGGC CGCTTGGGTG GAGAGGCTAT TCAGCTATGA CTGGGCACAA
 4441 CAGACAATCG GCTGCTCTGA TGCCGCCGTG TTCCGGCTGT CAGCGCAGGG GCGCCCGGTT
 4501 CTTTTGTCA AGACCGACCT GTCCGGTGCCTA CTGAATGAAC TGCAGGACGA GGCAGCCGG
 4561 CTATCGGGC TGGCACGAC GGGCGTTCT TGCGCAGCTG TGCTCGACGT TGTCACTGAA
 4621 GCGGGAAAGGG ACTGGCTGCT ATTGGGCGAA GTGCCGGGGC AGGATCTCCT GTCATCTCAC
 4681 CTTGCTCTG CCGAGAAAGT ATCCATCATG GCTGATGCAA TGCGCCGGCT GCATACGCTT
 4741 GATCCGGCTA CCTGCCATT CGACCACAA GCGAACACATC GCATCGAGCG AGCACGTACT
 4801 CGGATGGAAG CCGGTCTTGT CGATCAGGAT GATCTGGACG AAAGACATCA GGGGCTCGCG
 4861 CCAGCCGAAC TGTCGCCAG GCTCAAGGCC CGCATGCCCG ACGGCGAGGA TCTCGCTGTG
 4921 ACCCATGGCG ATGCCCTGCTT GCCGAATATC ATGGTGGAAA ATGGCCGCTT TTCTGGATT
 4981 ATCGACTGTG GCCGGCTGGG TGTGGCGGAC CGCTATCAGG ACATAGCGTT GGCTACCCGT
 5041 GATATTGCTG AAGAGCTTGG CGGCGAATGG GCTGACCGCT TCCTCGTGCT TTACGGTATC
 5101 GCCGCTCCCG ATTGCGAGCG CATGCCCTTC TATGCCCTTC TTGACGAGTT CTTCTGAGCG
 5161 GGACTCTGGG GTTCGAAATG ACCGACCAAG CGACGCCAA CCTGCCATCA CGATGGCCGC
 5221 AATAAAATAT CTTTATTTC ATTACATCTG TGTGTGGTT TTTTGTGTGA ATCGATAGCG
 5281 ATAAGGATCC GCGTATGGTG CACTCTCACT ACAATCTGCT CTGATGCCGC ATAGTTAAC
 5341 CAGCCCCGAC ACCCGCCAAC ACCCGCTGAC GCGCCCTGAC GGGCTTGTCT GCTCCGGCA
 5401 TCCGCTTACA GACAAGCTGT GACCGCTCTC GGGAGCTGCA TGTGTCAAGAG GTTTTCACCG
 5461 TCATCACCGA AACCGCGAG AGCAGGGC CTCGTGATAC GCCTATTIT ATAGGTTAAT
 5521 GTCATGATAA TAATGGTTTC TTAGACGTCA GGTGGCACTT TTGGGGAA TGTCGCGGGA
 5581 ACCCCTATTG TTTTATTAA CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA
 5641 CCCTGATAAA TGCTTCAATA ATATTGAAA AGGAAGAGTA TGAGTATCA ACATTTCCGT
 5701 GTGCCCTTA TTCCCTTTTG TGCGGCATT TGCGCTCTG TTTTGTCTCA CCCAGAAACG
 5761 CTGGTGAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACAG
 5821 GATCTCAACA GCGGTAAGAT CCTTGAGAGT TTTCGCCCG AAGAACGTTT TCCAATGATG
 5881 AGCACTTTA AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG-

FIGURE 32C

5941 CAACTCGGTC GCCGCATAACA CTATTCTAG AATGACTTGG TTGAGTACTC ACCAGTCACA
6001 GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG
6061 AGTGATAACA CTGGGGCCAA CTTACTTCTG ACAACCGATCG GAGGGACCGAA GGAGCTAAC
6121 GCTTTTTTGC ACAACATGGG GGATCATGTA ACTCGCCCTTG ATCGTTGGGA ACCGGAGCTG
6181 AATGAAGCCA TACCAAACGA CGAGCGTGAC ACCACCGATGC CTGTAGCAAT GGCAACAAACG
6241 TTGCGCAAAAC TATTAACTGG CGAAGTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC
6301 TGGATGGAGG CGGATAAAAGT TGCAAGGACCA CTTCTCGCCT CGGCCCTTCC GGCTGGCTGG
6361 TTTATTGCTG ATAAATCTGG AGCCGGTAGG CGTGGGTCTC GCGGTATCAT TGCAGCACTG
6421 GGGCCAGATG GTAAGCCCTC CGCTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT
6481 ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA
6541 CTGTCAGACC AAGTTTACTC ATATATACCT TAGATTGATT TAAAAACTTCA TTTTTAATTT
6601 AAAAGGATCT AGGTGAAGAT CCTTTTTGAT AATCTCATGA CAAAAATCCC TTAACGTGAG
6661 TTTTCGTTCC ACTGAGCGTC AGACCCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCCT
6721 TTTTTTCTGC GCGTAATCTG CTGCTTGCCT ACACAAAAAC CACCGCTTAC AGCGGTGGTT
6781 TGTTTGGCCGG ATCAAGAGCT ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG
6841 CAGATACCAA ATACTGTCCT TCTAGTGTAG CCGTAGTTAG GCCACCCTT CAAGAACTCT
6901 GTAGCACCGC CTACATACCT CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC
6961 GATAAGTCGT GTCTTACCGG GTTGGACTCA AGACGATAGT TACCGGATAA GGCGCAGCGG
7021 TCGGGCTGAA CGGGGGGTTC GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA
7081 CTGAGATACC TACAGCGTGA GCATTGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG
7141 GACAGGTATC CGGTAAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGG GCTTCCAGGG
7201 GGAAACGCCCT GGTATCTTAA TAGTCCTGTC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA
7261 TTTTTGTGAT GCTCGTCA

FIGURE 32D

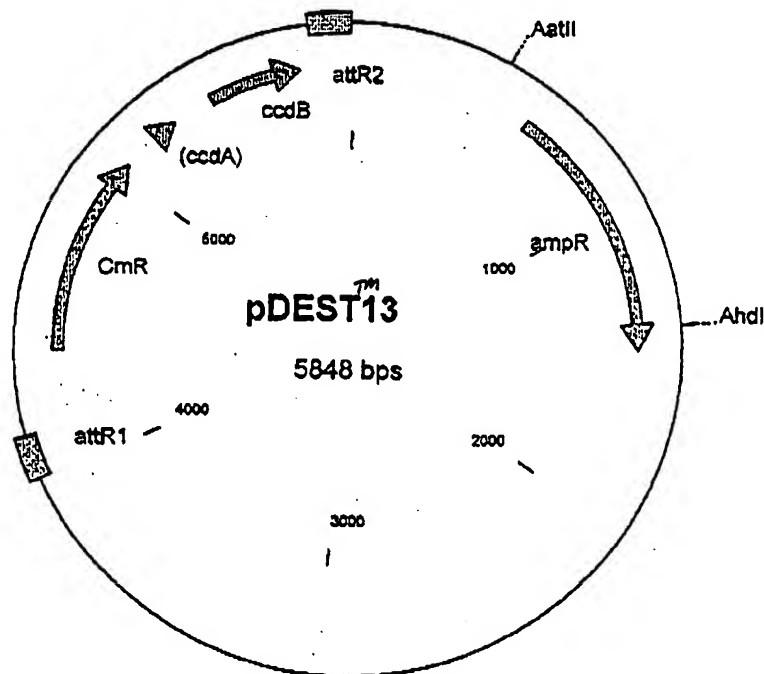
80/260

Figure 33A: pDEST13

Native protein in E. coli: λPL
promoter

BglII

3721 tgggcaaacc aagacagcta aagatctctc acctacccaa caatgc~~cccc~~ ctgcaaaaaa
 acccg~~ttt~~gg ttctgtcgat ttcttagagag tggatgg~~ttt~~ gttacgggg gacgttttt
 3781 taaaattcata taaaaaacat acagataacc atctgcggtg ataaattatc tctggcggtg
 attaaagtat atttttgtt tg~~tctat~~ttgg tagacgcccac tatttaatag agaccgcccac
 3841 -35 λPL Promoter -10 mRNA
 ttgacataaa taccactggc ggtgatactg agcacatcg caggacgcac tgaccaccat
 aactgtat~~ttt~~ atgg~~t~~gaccg ccactatgac tcgtgtagtc gcctcgctg actgg~~t~~ggta
 3901 gaaggtgacg ctcttaaaaa ttaageccctg aaaaaaggca gcattcaaaag cagaaggctt
 cttccactgc gagaattttt aatcgggac ttttccgt cgtaagttc gtttccgaa
 3961 tggggtgtgt gatacgaaac gaagcattgg gatcatcaca agtttgtaca aaaaagctga
 accccacaca ctatgcttg ct~~c~~tgtaacc ct~~c~~tagtgt tcaa~~c~~atgt ttttcgact,
 ↓



81/240

pDEST13 5848 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
599..1458	ampR
4123..3998	attR1
4372..5031	CmR
5151..5235	inactivated ccdA
5373..5678	ccdB
5719..5843	attR2

1 TTCACTGGCC GTCGTTTAC AACGTCGTGA CTGGAAAAAC CCTGGCGTTA CCCAACTTAA
 61 TCGCTTGCA GCACATCCCC CTTTCGCCAG CTGGCGTAAT AGCGAAGAGG CCCGCACCGA
 121 TCGCCCTTCC CAACAGTTGC GCAGCCTGAA TGGCGAATGG CGCCTGATGC GGTATTTCT
 181 CCTTACGCAT CTGTGCGTA TTTCACACCG CATATGGTGC ACTCTCGATA CAATCTGCTC
 241 TGATGCCGCA TAGTTAACGC AGCCCCGACA CCCGCCAACA CCCGCTGACG CGCCCTGACG
 301 GGCTTGTCTG CTCCCACGAT CCGCTTACAG ACAAGCTGTG ACCTGCTCCG GGAGCTGCAT
 361 GTGTCAGAGG TTTTCACCGT CATCACCGAA ACGGCGAGA CGAAAGGGC TCGTGATACG
 421 CCTATTTTA TAGGTTAATG TCATGATAAT AATGGTTCT TAGACGTCA GTGGCACTTT
 481 TCGGGAAAT GTGCGCGAA CCCCTATTG TTATTTTTC TAAATACATT CAAATATGTA
 541 TCCGCTCATG AGACAATAAC CCTGATAAAAT GCTTCAATAA TATTGAAAAA GGAAGAGTAT
 601 GAGTATTCAA CATTTCCTGT TCGCCCTTAT TCCCTTTTT GCGGCATTTT GCCTTCCTGT
 661 TTTTGCTCAC CCAGAAACGC TGTTGAAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG
 721 AGTGGTTAC ATCGAACTGG ATCTCAACAG CGGTAAAGATC CTTGAGAGTT TTCGCCCCGA
 781 AGAACGTTT CCAATGATGA GCACTTTAA AGTTCTGCTA TGTGGCGGG TATTATCCCG
 841 TATTGACGCC GGGCAAGAGC AACTCGTGTG CGCAGATACAC TATTCTCAGA ATGACTTGGT
 901 TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG
 961 CAGTGCTGCC ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTCTGA CAACGATCGG
 1021 AGGACCGAAG GAGCTAACCG CTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCTTGA
 1081 TCGTTGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC
 1141 TGTAGCAATG GCAACAACGT TGCGAAACT ATTAACCTGGC GAAACTACTTA CTCTAGCTTC
 1201 CGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAAGTT GCAGGACAC TTCTGCGCTC
 1261 GGCCCTTCCG GCTGGCTGG TTATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG
 1321 CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC
 1381 GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC
 1441 ACTGATTAAG CATTGGTAAC TGTCAAGACCA AGTTTACTCA TATATACTTT AGATTGATT
 1501 AAAACTTCAT TTTAATTAA AAAGGATCTA GGTGAAGATC CTTTTGATA ATCTCATGAC
 1561 CAAAATCCCT TAACGTGAGT TITCGTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA
 1621 AGGATCTTCT TGAGATCCTT TTTTCTGCG CGTAATCTGC TGCTTGCAA CAAAAAAACC
 1681 ACCGCTACCA GCGGTGGTTT GTTTGCCTGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT
 1741 AACTGGCTTC AGCAGAGCGC AGATACAAA TACTGTTCTT CTAGTGTAGC CGTAGTTAGG
 1801 CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCCTGTTACC
 1861 AGTGGCTGCT GCCAGTGGCG ATAAGCTGTG TCTTACCGGG TTGGACTCAA GACGATAGTT
 1921 ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGG
 1981 GCGAACGACC TACACCGAAC TGAGATAACCT ACAGCGTGAG CATTGAGAAA GCGCACGCT
 2041 TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAAGCGGC AGGGTCGGAA CAGGAGAGCG
 2101 CACCGAGGGAG CTTCCAGGGG GAAACGCCCTG GTATCTTAT AGTCTGTGCG GGTTCGCCA
 2161 CCTCTGACTT GAGCGTGTAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA
 2221 CGCCAGCAAC GCGGCCTTTT TACGGTTCTT GGCCTTTGTC TGGCTTTTG CTCACATGTT
 2281 CTTTCCTGCG TTATCCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA
 2341 TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCCGAAGA
 2401 GCGCCCAATA CGCAAACCGC CTCTCCCCGC CGCTTGGCCG ATTCTTAAT GCAGCTGGCA
 2461 CGACAGGTTC CCCGACTGGA AAGCGGGAG TGAGCGAAC GCAATTAAAT TGAGTTAGCT
 2521 CACTCATTAG GCACCCAGG CTTTACACTT TATGCTTCCG GCTCGTATGT TGTGTGGAAT
 2581 TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC CATGATTACG CCAAGCTTGG
 2641 CTGCAGGTGA TGATTATCAG CCAGCAGAGA TTAAGGAAAA CAGACAGGTT TATTGAGCGC
 2701 TTATCTTCC CTTTATTTT GCTGCGGTAA GTGCGATAAA AACCATTCTT CATAATTCAA-

FIGURE 33B

82/240

2761 TCCATTACT ATGTTATGTT CTGAGGGGAG TGAAAATTCC CCTAATTCGA TGAAGATTCT
 2821 TGCTCAATTG TTATCAGCTA TGCGCCGACC AGAACACCTT GCCGATCAGC CAAACGTCTC
 2881 TTCAGGCCAC TGACTAGCGA TAACCTTCCC CACAACGGAA CAACTCTCAT TGCATGGAT
 2941 CATTGGGTAC TGTGGGTTTA GTGGTTGAA AAACACCTGA CGCGTATCCC TGATCAGTT
 3001 CTTGAAGGTA AACTCATCAC CCCCAAGTCT GGCTATGCAG AAATCACCTG GCTAACACG
 3061 CTGCTCAGGG TCAACGAGAA TTAACATTCC GTCAGGAAAG CTTGGCTTGG AGCCTGTTGG
 3121 TGCGGTCTAT GAATTACCTT CAACCTCAAG CCAGAATGCA GAATCACTGG CTTTTTGTT
 3181 TGTGCTTACCATCTCAGGGTAA AACACAGCTA CTCATACTCA CTTCTAAGTG ACGGCTGCAT
 3241 CCCTGCCTGA ACATGAGAAA CGTAGATTTC TCTGGCGATT GAAGGGCTAA ATTCTTCAC
 3301 ACTAACCGCT TCATACATCT CAAGCAATGC GGCCTTATAA GCATTTAATG CATTGATGCC
 3361 GCTAACTTTG AGAATTTTTG CTGACTGCC CATCCCCATC TTGTCTGCAG CAGATTCTG
 3421 ATTAAATAAA GCACCAACGC TCTTTTTTC ATAAATTGCT TTAAGGCAG GTGCGTCCTC
 3481 GGATAAGCCA AGTTCATTTT GTTTCTTTT TGTGCTCATA CGTTAAATCT ATCACCGCAA
 3541 AAGCTGCTCT TGTGTTAATG TGCGTGTGA CTATTTTAC TCTGGCGGTG ATAATGGTTG
 3601 GGGATAATAA TCTAACACCG TGGAAACAAG CATAACCCCTG AAAGATTATG CAATGCCTT
 3661 CATGTACTAA GGAGGTTGTA AAGATCTCTC ACCTACAAA CAATGCCCTC CTGCAAAAAA
 3721 TGGGCAAACC AAGACAGCTA ACAGATAACC ATCTGCGGTG ATAAATTATC TCTGGCGGTG
 3781 TAAATTCTATA TAAAAAACAT 3841 TTGACATAAA TACCACTGGC GGTGATACTG AGCACATCAG CAGGACGCAC TGACCACCAT
 3901 GAAGGTGACG CTCTTTAAAAA TTAAGCCCTG AAGAAGGGCA GCATTCAAAG CAGAAGGCTT
 3961 TGGGGTGTGT GATACGAAAC GAAGCATTGG GATCATCACA AGTTTGTACA AAAAAGCTGA
 4021 ACGAGAAACG TAAAATGATA TAAATATCAA TATATTAAAT TAGATTTGC ATAAAAAAACA
 4081 GACTACATAA TACTGTAAAA CACAACATAT CCAGTCACTA TGGCGGCCGC TAAGTTGGCA
 4141 GCATCACCCG ACCGACTTTG CGCCGAATAA ATACCTGTGA CGGAAGATCA CTTCGCAGAA
 4201 TAAATAAATC CTGGTGTCCC TGTTGATACC GGGAAAGCCCT GGGCCAACCTT TTGGCGAAAA
 4261 TGAGACGTTG ATCGGCACGT AAGAGTTCC AACTTCACC ATAATGAAAT AAGATCACTA
 4321 CGGGCGTAT TTTTGAGTT ATCGAGATT TCAGGAGCTA AGGAAGCTAA ATGGAGAAA
 4381 AAAATCACTG GATATACCAC CGTTGATATA TCCCATGGC ATCGTAAAGA ACATTTGAG
 4441 GCATTTCAGT CAGTTGCTCA ATGTACCTAT AACCAAGACCG TTCAGCTGGA TATTACGGCC
 4501 TTTTTAAAGA CCGTAAAGAA AAATAAGCAC AAGTTTATC CGGCCTTAT TCACATTCTT
 4561 GCCCCCTGA TGAATGCTCA TCCGGAATTC CGTATGGCAA TGAAAGACGG TGAGCTGGTG
 4621 ATATGGATA GTGTTCACCC TTGTTACACC GTTTCATCATG AGCAAACCTGA AACGTTTCA
 4681 TCGCTCTGGA GTGAATACCA CGACGATTTC CGGCAGTTTC TACACATATA TTCGCAAGAT
 4741 GTGGCGTGT ACAGGTGAAAA CCTGGCCTAT TTCCCTAAAG GTTTTATTGA GAATATGTTT
 4801 TTCGCTCTCAG CCAATCCCTG GGTGAGTTTC ACCAGTTTG ATTIAAACGT GGCCAAATATG
 4861 GACAACTTCT TCGCCCCCGT TTTCACCATG GGCAAATATT ATACGCAAGG CGACAAGGTG
 4921 CTGATGCCGC TGGCGATTCA GTTTCATCAT GCCGCTCTG ATGGCTTCCA TGTCGGCAGA
 4981 ATGCTTAATG AATTACAACA GTACTGCGAT GAGTGGCAGG GCGGGGGCGTA AACGCGTGG
 5041 TCCGGCTTAC TAAAAGCCAG ATAACAGTAT GCGTATTGCG GCGCTGATT TTGCGGTATA
 5101 AGAATATATA CTGATATGTA TACCCGAAGT ATGCAAAAAA GAGGTGTGCT ATGAAGCAGC
 5161 GTATTACAGT GACAGTTGAC AGCGACAGCT ATCAGTTGCT CAAGGATATA ATGATGTCAA
 5221 TATCTCCGGT CTGGTAAGCA CAACCATGCA GAATGAAGCC CGTCGCTGCG GTGCCGAACG
 5281 CTGGAAAGCG GAAAATCAGG AAGGGATGGC TGAGGTCGCC CGGTTTATTG AAATGAACGG
 5341 CTCTTTGCT GACGAGAACAA GGGACTGGT AAATGCAGTT TAAGGTTTAC ACCTATAAAA
 5401 GAGAGAGCCG TTATCGTCTG TTGTTGATG TACAGAGTGA TATTATTGAC ACGCCCGGGC
 5461 GACGGATGGT GATCCCCCTG GCCAGTGCAC GTCTGCTGTC AGATAAAAGTC TCCCGTGAAC
 5521 TTGACCCGGT GGTGCATATC GGGGATGAAA GCTGGCGCAT GATGACCACC GATATGGCCA
 5581 GTGTGCCGGT CTCCGTTATC GGGGAAGAAG TGGCTGATCT CAGCCACCGC GAAAATGACA
 5641 TCAAAAACGC CATTAAACCTG ATGTTCTGGG GAATATAAT GTCAGGCTCC GTTATACACA
 5701 GCCAGTCTGC AGGTGCGACCA TAGTGACTGG ATATGTTGTG TTGTTACAGTA TTATGTAGTC
 5761 TGTTTTTAT GCAAAATCTA ATTAAATATA TTGATATTAA TATCAATTAA CGTTTCTCGT
 5821 TCAGCTTCT TGTACAAAGT GGTGATAA

FIGURE 33C

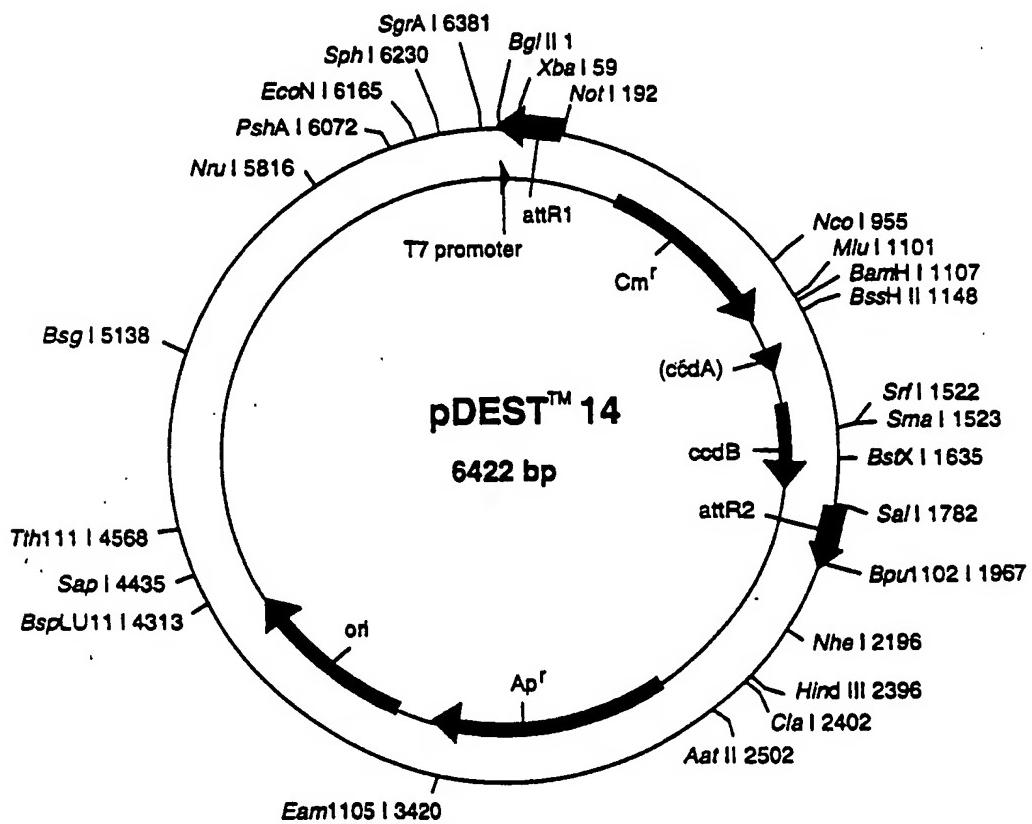
83/240

Figure 34A: pDEST14 Native Protein Expression in *E. coli*, T7 Promoter

3961 tgccggccac gatgcgtccg gcgttagagga tcgatatctc gatcccgcga aataataacg
 acggccggtg ctacgcaggc cgcatttcct agctcttag ctagggcgct ttatttatatgc
 M_PVA

4021 // actcacata gggagaccac aacggtttcc ctcttagatca caagttttgt caaaaaagct
 tgagtatat cccctctggtg ttgccaagg gagatcttagt gttcaaacat gtttttcga //

BglII AclI PT7 →
 XbaI attR1



84/240

pDEST14 6422 bp (rotated to position 4000)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
185..61	attR1
435..1094	CmR
1214..1298	inactivated ccda
1436..1741	ccdB
1782..1906	attR2
2632..3489	ampR
1 CGATCCCGCG AAATTAATAC GACTCACTAT AGGGAGACCA CAACGGTTTC CCTCTAGATC	
61 ACAAGTTGT ACAAAAAAGC TGAACGAGAA ACGTAAAATG ATATAAATAT CAATATATTA	
121 AATTAGATTT TGCAAAAAA ACAGACTACA TAATACTGTA AAACACAACA TATCCAGTCA	
181 CTATGGCGGC CGCTAACGTT GCAGCATCAC CCGACGCACT TTGCGCCGAA TAAATACCTG	
241 TGACGGAAGA TCACTTCGCA GAATAAAATA ATCCTGGTGT CCCTGTTGAT ACCGGGAAGC	
301 CCTGGGCCAA CTTTGGCGA AAATGAGACG TTGATCGGCA CGTAAGAGGT TCCAACCTTC	
361 ACCATAATGA AATAAGATCA CTACCGGGCG TATTTTTGAGA GTTATCGAGA TTTTCAGGAG	
421 CTAAGGAAGC TAAAATGGAG AAAAAAAATCA CTGGATATAC CACCGTTGAT ATATCCAAT	
481 GGCATCGTAA AGAACATTTT GAGGCATTTT AGTCAGTTGC TCAATGTACC TATAACCAGA	
541 CCGTTCAGCT GGATATTACG GCCTTTTAA AGACCGTAAA GAAAAATAAG CACAAGTTT	
601 ATCCGGCCTT TATTACACATT CTTGCCCCGC TGATGAATGC TCATCCGGAA TTCCGTATGG	
661 CAATGAAAGA CGGTGAGCTG GTGATATGGG ATAGTGTCA CCCTTGTAC ACCGTTTCC	
721 ATGAGCAAAC TGAAACGTT TCATCGCTCT GGAGTGAATA CCACGACGAT TTCCGGCAGT	
781 TTCTACACAT ATATTGCAA GATGTGGCGT GTTACGGTGA AACACCTGGCC TATTTCCCTA	
841 AAGGGTTTAT TGAGAATATG TTTTCGCTC CAGCCAATCC CTGGGTGAGT TTCACCAGTT	
901 TTGATTTAAA CGTGGCCAAT ATGGACAATCT TCTTCGCCCC CGTTTTCACC ATGGGCAAAT	
961 ATTATACGCA AGGCAGACAAG GTGCTGATGC CGCTGGCGAT TCAGGTTCAT CATGCCGTCT	
1021 GTGATGGCTT CCATGTCGGC AGAATGCTTA ATGAATTACA ACAGTACTGC GATGAGTGGC	
1081 AGGGCGGGGC GTAAACGCGT GGATCCGGCT TACTAAAAGC CAGATAACAG TATGCGTATT	
1141 TCGCGCTGA TTTTGCGGT ATAAGAATAT ATACTGATAT GTATACCCGA AGTATGTCAA	
1201 AAAGAGGTGT GCTATGAAGC AGCGTATTAC AGTGCACAGTT GACAGCGACA GCTATCAGTT	
1261 GCTCAAGGCA TATATGATGT CAATATCTCC GGTCTGGTAA GCACAACCAT GCAGAATGAA	
1321 GCCCCGTCGTC TCGGTGCCGA ACGCTGGAAA CGGGAAAATC AGGAAGGGAT GGCTGAGGTC	
1381 GCCCCGTTTA TTGAAATGAA CGGCTCTTT GCTGACGAGA ACAGGGACTG GTGAAATGCA	
1441 GTTTAAGGTT TACACCTATA AAAGAGAGAG CCGTTATCGT CTGTTTGTGG ATGTACAGAG	
1501 TGATATTATT GACACGGCCCG GGCGACGGAT GGTGATCCCC CTGGCCAGTG CACGTCTGCT	
1561 GTCAGATAAA GTCTCCCGTG AACITTAACCC GGTGGTGCAT ATCGGGGATG AAAGCTGGCG	
1621 CATGATGACC ACCGATATGG CCAGTGTGCC GGTCTCCGTT ATCGGGGAAAG AAGTGGCTGA	
1681 TCTCAGGCCAC CGCGAAAATG ACATCAAAAA CGCCATTAAC CTGATGTTCT GGGGAATATA	
1741 AATGTCAGGC TCCCTTATAC ACAGCCAGTC TGCAGGTGCA CCATAGTGAC TGGATATGTT	
1801 GTGTTTACA GTATTATGTA GTCTGTTTT TATGAAAAT CTAATTTAAT ATATTGATAT	
1861 TTATATCATT TTACGTTTCT CGTTCAGCTT TCTTGACAA AGTGGTGATG ATCCGGCTGC	
1921 TAACAAAGCC CGAAAGGAAG CTGAGTTGGC TGCTGCCACC GCTGAGCAAT AACTAGCATA	
1981 ACCCCCTGGG GCCTCTAACCGGGG GGGTCTTGAG GGGTTTTTG CTGAAAGGAG GAACTATATC	
2041 CGGATATCCA CAGGACGGGT GTGGTCGCCA TGATCGCGTA GTCGATAGTG GCTCCAAGTA	
2101 GCGAAGCGAG CAGGACTGGG CGGCGGCCAA AGCGCTCGGA CAGTGCTCCG AGAACGGGTG	
2161 CGCATAGAAA TTGCAATCAC GCATATAGCG CTAGCAGCAC GCCATAGTGA CTGGCGATGC	
2221 TGTCGGAATG GACGATATCC CGCAAGAGGC CGGGCAGTAC CGGCATAACC AAGCCTATGC	
2281 CTACAGCATC CAGGGTGACG GTGCCGAGGA TGACGATGAG CGCATTGTTA GATTTCATAC	
2341 ACGGTGCGCTG ACTGGCTTAG CAATTTAACT GTGATAAAACT ACCGCATTAA AGCTTATCGA	
2401 TGATAAGCTG TCAAACATGA GAATTCTGAA AGACGAAAGG GCCTCGTGAT ACGCCTATT	
2461 TTATAGGTTA ATGTCATGAT AATAATGGTT TCTTAGACGT CAGGTGGCAC TTTTCGGGGA	
2521 AATGTCGCGCG GAACCCCTAT TTGTTTATT TTCTAAATAC ATTCAAATAT GTATCCGCTC	
2581 ATGAGACAAT AACCCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT	
2641 CAACATTCC GTGTCGCCCT TATTCCCTT TTTGCGGCAT TTGCGCTTCC TGTTTTGCT	
2701 CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT-	

FIGURE 34B

85/240

2761 TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTCGCC CGAAGAACGT
 2821 TTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTGTGAC
 2881 GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC
 2941 TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT
 3001 GCCATAACCA TGAGTGTATAA CACTGCGGCC AACTIACCTC TGACAACGAT CGGAGGACCG
 3061 AAGGAGCTAA CCGCTTTTT GCACAAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG
 3121 GAACCGGAGC TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCAACGAT GCCTGCAGCA
 3181 ATGGCAACAA CGTTGCGCAA ACTATTAACG GGCAGAACTAC TTACTCTAGC TTCCCGGCAA
 3241 CAATTAATAG ACTGGATGGA GCGGGATAAA GTTGCAGGAC CACTTCTGCG CTCGGCCCTT
 3301 CCGGCTGGCT GTTTTATTGC TGATAAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC
 3361 ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGGG
 3421 AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT
 3481 AAGCATTGGT AACTGTCAGA CCAAGTTTAC TCATATATAC TTAGATTGA TTTAAAACCTT
 3541 CATTTTTAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTG ATAATCTCAT GACCAAAATC
 3601 CCTTAACGTG AGTTTCGTT CCACTGAGCG TCAGACCCCC TAGAAAAAGAT CAAAGGATCT
 3661 TCTTGAGATC CTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA
 3721 CCAGCGGTGG TTTGTTGCC GGATCAAGAG CTACCAACTC TTTTTCCGAA GGTAACTGGC
 3781 TTCAGCAGAG CGCAGATAACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACCCAC
 3841 TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCTGTT ACCAGTGGCT
 3901 GCTGCCAGTG GCGATAAGTC GTGTCTTAC GGGTTGGACT CAAGACGATA GTTACGGGAT
 3961 AAGGCCAGC GGTGGGCTG AACGGGGGGT TCGTCACAC AGCCCGAGCTT GGAGCGAACG
 4021 ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA
 4081 GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCAGGAGG
 4141 GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA
 4201 CTTGAGCGTC GATTTTTGTG ATGTCGTCA GGGGGCGGA GCCTATGGAA AAACGCCAGC
 4261 AACGCCGCT TTTTACGGTT CCTGGCCTT TGCTGGCCTT TTGCTCACAT GTTCTTTCT
 4321 GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAAGC TGATACCGCT
 4381 CGCCGAGGCC GAACGACCGA GCGCAGCGAG TCAGTGAAGC AGGAAGCGGA AGAGCGCCTG
 4441 ATGCGGTATT TTCTCCCTAC GCATCTGTG GGTATTTCAC ACCGCATATA TGGTGCACTC
 4501 TCAGTACAAT CTGCTCTGAT GCGCGATAGT TAAGCCAGTA TACACTCCGC TATCGCTACG
 4561 TGACTGGGTC ATGGCTGC CGCGACACCC GCCAACACCC GCTGACGGGC CCTGACGGGC
 4621 TTGTCGTCTC CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA GCTGCATGTG
 4681 TCAGAGGTTT TCACCGTCA CACCGAAAC CGCGAGGAGC CTGCGTAA GCTCATCAGC
 4741 GTGGTCGTGA AGCGATTCAAC AGATGTCTGC CTGTCATCC GCTTCAGCT CGTTGAGTTT
 4801 CTCCAGAACG GTTAATGTCT GGCTTCTGAT AAAGCGGGCC ATGTTAAGGG CGGTTTTTC
 4861 CTGTTGGTC ACTGATGCCT CCGTGAAGG GGGATTTCTG TTGATGGGG TAATGATACC
 4921 GATGAAACGA GAGAGGATGC TCACGATACG GGTTACTGAT GATGAACATG CCCGGTTACT
 4981 GGAACGTTGT GAGGGTAAAC AACTGGGGT ATGGATGCC CGGGGACAGA GAAAATCAC
 5041 TCAGGGTCAA TGCCAGCGCT TCGTTAACAC AGATGTAGGT GTTCCACAGG GTAGCCAGCA
 5101 GCATCTGCG ATGCAGATCC GGAACATAAT GGTGCAGGGC GCTGACTTCC GCGTTTCCAG
 5161 ACTTTACGAA ACACGGAAAC CGAAGACCAT TCATGTTGTT GCTCAGGTGC CAGACGTTT
 5221 GCAGCAGCAG TCGCTTCACG TTGCGCTCGCG TATCGGTGAT TCATTCTGCT AACCAGTAAG
 5281 GCAACCCCGC CAGCCTAGCC GGGTCTCAA CGACAGGAGC ACGATCATGC GCACCCGTGG
 5341 CCAGGAGCCA AGCGCTGGCG AGATGCGCCG CGTGCAGGCTG CTGGAGATGG CGGACCGCAT
 5401 GGATATGTTG TGCCAAAGGGT TGGTTGGCG ATTCAACAGTT CTCCGCAAGA ATTGATTGGC
 5461 TCCAATTCTT GGAGTGGTGA ATCCGTTAGC GAGGTGCCGC CGGTTCCAT TCAGGTCGAG
 5521 GTGGCCCGGC TCCATGCACC GCGACCAAC GCGGGGAGGC AGACAAGGTA TAGGGCGCG
 5581 CCTACAATCC ATGCCAACCC GTTCCATGTG CTCCGGAGG CGGCATAAAAT CGCCGTGACG
 5641 ATCAGCGGTC CAGTGTGATCA AGTTAGGTG GTAAGAGCCG CGAGCGATCC TTGAAGCTGT
 5701 CCCTGATGGT CGTCATCTAC CTGCGCTGAC AGCATGGCCT GCAACGCCGG CATCCCGATG
 5761 CGGCCGGAAG CGAGAAGAAT CATAATGGGG AAGGCCATCC AGCCTCGCGT CGCGAACGCC
 5821 AGCAAGACGT AGCCCAAGCGC GTCGGCCGCC ATGCCGGCA TAATGGCCTG CTTCTCGCG
 5881 AAACGTTGG TGCGGGAGC AGTGAAGAAG GCTTGAGCGA GGGCGTGCAA GATTCCGAAT
 5941 ACCGCAAGCG ACAGGCCGAT CATCGTCGGC CTCCAGCGA AGCGGTCTC GCCGAAATG
 6001 ACCCAGAGCG CTGCGGGCAC CTGTCCTACG AGTTGCATGA TAAAGAAGAC AGTCATAAGT
 6061 GCGCGACGA TAGTCATGCC CCGCGCCAC CGGAAGGAGC TGACTGGGTT GAAGGCTCTC
 6121 AAGGGCATCG GTCGATCGAC GCTCTCCCTT ATGCGACTCC TGCATTAGGA AGCAGCCAG
 6181 TAGTAGGTTG AGGCCGTGA GCACCGCCGC CGCAAGGAAT GGTGCATGCA AGGAGATGGC-

FIGURE 34C

86/240

6241 GCCCAACAGT CCCCCGGCCA CGGGGCCTGC CACCATAACCC ACGCCGAAAC AAGCGCTCAT
6301 GAGCCCGAAG TGGCGAGCCC GATCTTCCCC ATCGGTGATG TCGGCGATAT AGGCCGCCAGC
6361 AACCGCACCT GTGGCGCCGG TGATGCCGGC CACCGATGCGT CCGGCGTAGA GGATCGAGAT
6421 CT

FIGURE 34D

87/240

Figure 35A: pDEST15 Glutathione-S-transferase Fusion in *E. coli*, T7 Promoter

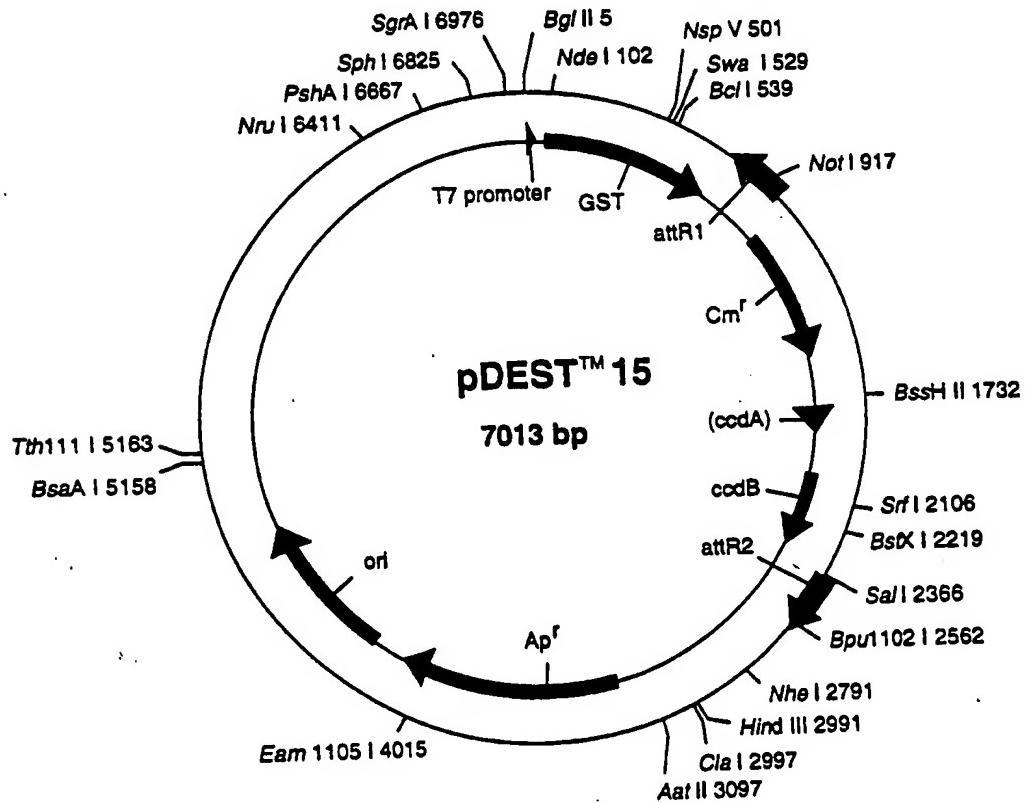
mRNA

T7 Promoter

```

1  nat cga gat ctc gat ccc gcg aaa qta ata cga ctc act ata [ggg] aga cca
    nta gct cta gag cta ggg cgc ttt qat tat get gag tga tat ccc tct ggt
      XbaI
52  caa cgg ttt ccc qet aga aat aat ttt gtt taa ctt taa gaa gga gat ata
    gtt gcc aaa ggg aga tct tta tta aaa caa att gaa att ctt cct cta tat
103 NdeI S P I L
    cat atg tcc cct aata cta ggt tat tgg aaa att aag ggc ctt gtg caa ccc
    gta tac agg gga tat gat cca ata acc ttt taa ttc ccg gaa cac gtt ggg
      Start Translaction GST
154  act cga ctt ctt ttg gaa tat ctt gaa gaa aaa tat gaa gag 'cat ttg tat
    tga gct gaa gaa aac ctt ata gaa ctt ctt ttt ata ctt ctc gta aac ata

715  cag ggc tgg caa gcc acg ttt ggt ggc gac cat cct cca aaa tcg gat
    gtc ccc acc gtt cgg tgc aaa cca cca ccc ctg gta gga ggt ttt agc cta
      S N Q T S L Y K K A
766  ctg gtt ccc cgt eca tgg tgg dat caa aca agt ttg tac aaa aaa gct gaa
    gac caa ggc gca ggt acc agc tta gtt tgg tca aac atg EEE ttt cga ctt
      attR1 attY
817  cga gaa acg taa aat gat ata aat atc aat ata tta aat tag att ttg cat
    gct ctt tgc att tta cta tat tta tag tta tat aat tta attc taa aac gta
  
```



88/240

pDEST15 7013 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
108..776	GST
916..792	attR1
1025..1537	CmR
1804..1888	inactivated ccdA
2026..2331	ccdB
2372..2496	attR2
3233..4093	ampR

1 ATCGAGATCT CGATCCCGCG AAATTAAATAC GACTCACTAT AGGGAGACCA CAACGGTTTC
 61 CCTCTAGAAA TAATTTTGTG TAACTTTAAG AAGGAGATAT ACATATGTCC CCTATACTAG
 121 GTTATTGGAA AATTAAAGGC CTTGTGCAAC CCACCTCGACT TCTTTTGAA TATCTTGAG
 181 AAAAATATGA AGAGCATTG TATGAGCGCG ATGAAGGTGA TAAATGGCGA AACAAAAAGT
 241 TTGAATTGGG TTGGAGTTT CCCAATCTTC CTTATTATAT TGATGGTGAT GTTAAATTAA
 301 CACAGTCTAT GGCCATCATA CGTTATATAG CTGACAAGCA CAACATGTG GGTGGTTGTC
 361 CAAAAGAGCG TGCAAGAGATT TCAATGCTTG AAGGAGCGGT TTTGGATATT AGATACGGTG
 421 TTTCGAGAAT TGCAATATAGT AAAGACTTTG AAACCTCTAA AGTTGATTTT CTTAGCAAGC
 481 TACCTGAAAT GCTGAAAATG TTCGAAGATC GTTTATGTCA TAAAACATAT TTAAATGGTG
 541 ATCATGTAAC CCATCCTGAC TTCAATGTTGAT ATGACGCTCT TGATGTTGTT TTATACATGG
 601 ACCCAATGTG CCTGGATGCG TTCCCAAAT TAGTTGTTT TAAAAAACGT ATTGAAGCTA
 661 TCCCACAAAT TGATAAGTAC TTGAATCCA GCAAGTATAT AGCATGGCCT TTGCAGGGCT
 721 GGCAAGCCAC GTTGGTGGT GGCGACCATC CTCCAAAATC GGATCTGGTT CCGCGTCAT
 781 GGTGAAATCA AACAAAGTTG TACAAAAAAG CTGAAACGAGA AACGTTAAAT GATATAAATA
 841 TCAATATATT AAATTAGATT TTGCATAAAA AACAGACTAC ATAATACTGT AAAACACAAAC
 901 ATATCCAGTC ACTATGGCGG CCGCAATTAGG CACCCCCAGGC TTACACTTT ATGCTTCGG
 961 CTCGTATAAT GTGTGGATT TGAGTTAGGA TCCGTCGAGA TTTCAGGAG CTAAGGAAGC
 1021 TAAAATGGAG AAAAAAAATCA CTGGATATAC CACCGTTGAT ATATCCAAAT GGCACTGTA
 1081 AGAACATTTT GAGGCATTTC AGTCAGTTGC TCAATGTACC TATAACCAGA CCGTTCAAGCT
 1141 GGATATTACG GCCTTTTAA AGACCGTAAA GAAAATAAG CACAAGTTT ATCCGGCCTT
 1201 TATTCACATT CTTGCCCGCC TGATGAATGC TCATCCGGAA TTCCGTATGG CAATGAAAGA
 1261 CGGTGAGCTG GTGATATGGG ATAGTGTTCAC CCCTGTTAC ACCGTTTTC ACCGTTTTC
 1321 TGAAACGTTT TCATCGCTCT GGAGTGAATA CCACGACGAT TTCCGGCAGT TTCTACACAT
 1381 ATATTCCAA GATGTGGCGT GTTACGGTGA AAACCTGGCC TATTTCCCTA AAGGGTTTAT
 1441 TGAGAAATATG TTTTCGCTCT CAGCCAATCC CTGGGTGAGT TTACCCAGTT TTGATTTAAA
 1501 CGTGGCCAAT ATGGACAAC TCTTCGCCCC CGTTTCACC ATGGGCAAAT ATTATACGCA
 1561 AGGCACAAAG GTGCTGATGC CGCTGGCGAT TCAGGTTCAT CATGCCGTCT GTGATGGCTT
 1621 CCATGTCGGC AGAATGCTTA ATGAATTACA ACAGTACTGC GATGAGTGGC AGGGCGGGGC
 1681 GTAATCTAGA GGATCCGGCT TACTAAAAGC CAGATAACAG TATGCGTATT TGCAGCTGA
 1741 TTTTGCGGTT ATAAGAAATAT ATACTGATAT GTATACCCGA AGTATGTCAA AAAGAGGTGT
 1801 GCTATGAAGC AGCGTATTAC AGTGACAGTT GACAGCGACA GCTATCAGTT GCTCAAGGCA
 1861 TATATGATGT CAATATCTCC GGTCTGGTAA GCACAACCAT GCAGAAATGAA GCCCCTCGTC
 1921 TGCCTGCCGA ACGCTGGAAA GCGGAAATC AGGAAGGGAT GGTGAGGTC GCCCCTGTTA
 1981 TTGAAATGAA CGGCTTTTGT GCTGACGAGA ACAGGGACTG GTGAAATGCA GTTTAAGGTT
 2041 TACACCTATA AAAGAGAGAG CCGTTATCGT CTGTTGTTGG ATGTACAGAG TGATATTATT
 2101 GACACGCCCG GGCACGGAT GGTGATCCCC CTGGCCAGTG CACGTCTGCT GTCAGATAAA
 2161 GTCTCCCGTG AACTTTACCC GGTGGTGCAT ATCGGGGATG AAAGCTGGCG CATGATGACC
 2221 ACCGATATGG CCAAGTGTGCC GGTCTCCGGT ATCGGGGAAAG AAGTGGCTGA TCTCAGCCAC
 2281 CGCGAAAATG ACATCAAAAA CGCCTTAAAC CTGATGTTCT GGGGAATATA AATGTCAGGC
 2341 TCCCTTATAC ACAGCCAGTC TGCAAGTGC GA CCACAGTGAC TGATGTTGTT GTGTTTACA
 2401 GTATTATGTA GTCTGTTTT TATGCAAAAT CTAATTTAAT ATATGATAT TTATATCATT
 2461 TTACGTTCT CGTTCAAGCTT TCTTGTACAA AGTGGTTTGA TTGACCCGG GATCCGGCTG
 2521 CTAACAAAGC CGGAAAGGAA GCTGAGTTGG CTGCTGCCAC CGCTGAGCAA TAACTAGCAT
 2581 AACCCCTGGG GGCCTCTAAA CGGGTCTTGA GGGGTTTTT GCTGAAAGGA GGAACCTATAT
 2641 CGGGATATCC ACAGGACGGG TGTGGTCGCC ATGATCGCGT AGTCGATAGT GGCTCCAAGT-

FIGURE 35B

89/240

2701 AGCGAAGCGA GCAGGACTGG GCGGCAGGCCA AAGCGGTGG ACAGTGCTCC GAGAACGGGT
 2761 GCGCATAGAA ATTGCATCAA CGCATATAAC GCTAGCAGCA CGCCATAGTG ACTGGCGATG
 2821 CTGTCGGAAT GGACGATATC CCGCAAGAGG CCCGGCAGTA CCGGCATAAC CAAGCCTATG
 2881 CCTACAGCAT CCAGGGTGCAC GGTGCCGAGG ATGACGATGA GGCATTGTT AGATTCATA
 2941 CACGGTGCCT GACTGCCTTA GCAATTAAAC TGTGATAAAC TACCGCATTAA AAGCTTATCG
 3001 ATGATAAGCT GTCAAACATG AGAATTCTTG AAGACGAAAG GGCCTCGTGA TAGCCTATT
 3061 TTATAGGTT AATGTCATGA TAATAATGGT TTCTTAGACG TCAGGGGCA CTTTCCGGG
 3121 AAATGTGCGC GGAACCCCTA TTGTTTATT TTTCTAAATA CATTCAAATA TGATCCGCT
 3181 CATGAGACAA TAACCCTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT
 3241 TCAACATTTC CGTGTGCCCC TTATTCCTT TTTGCGGCA TTTTGCCTTC CTGTTTTG
 3301 TCACCCAGAA ACGCTGGTGA AAGTAAAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG
 3361 TTACATCGAA CTGGATCTCA ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAACG
 3421 TTTTCCAATG ATGAGCACTT TTAAAGTTCT GCTATGTGGC GCGGTATTAT CCCGTGTTGA
 3481 CGCCGGGCAA GAGCAACTCG GTCGCCGCAT ACACATTCT CAGAATGACT TGGTTGAGTA
 3541 CTCACCAAGTC ACAGAAAAGC ATCTTACGGG TGCGATGACA GTAAGAGAAT TATGCAGTGC
 3601 TGCCATAACC ATGAGTGATA ACACGTGGC CAACTTACTT CTGACAAACGA TCGGAGGACC
 3661 GAAGGAGCTA ACCGTTTTT TGCAACACAT GGGGGATCAT GTAACTCGCC TTGATCGTTG
 3721 GGAACCGGAG CTGAATGAAG CCATACCAAA CGACGAGCGT GACACCACGA TGCCTGCAGC
 3781 AATGGCAACA ACGTTGCGCA AACTATTAAAC TGGCGAACTA CTTACTCTAG CTTCCGGCA
 3841 ACAATTAAATA GACTGGATGG AGGCGGATAA AGTTGCGAGGA CCACTTCTGC GCTCGGCCCT
 3901 TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT GAGCGTGGGT CTCGCGGTAT
 3961 CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT ACACGACGGG
 4021 GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
 4081 TAAGCATTGG TAACGTGTCAG ACCAAGTTA CTCATATATA CTTTAGATTG ATTTAAAAT
 4141 TCATTTTAA TTTAAAAGGA TCTAGGTGA GATCCTTTT GATAATCTCA TGACCAAAAT
 4201 CCCTTAACGT GAGTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC
 4261 TTCTTGAGAT CCTTTTTTC TGCGCGTAAT CTGCTGTTG CAAACAAAAA AACCACCGCT
 4321 ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTCCGA AGGTAACTGG
 4381 CTTCAGCAGA GCGCAGATAAC CAAATACTGT CCTCTAGTG TAGCCGTAGT TAGGCCACCA
 4441 CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAAGTGGC
 4501 TGCTGCCAGT GGCATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA
 4561 TAAGGCGCAG CGGTGGGCT GAACGGGGGG TTGCTGCACA CAGCCCAGCT TGGAGCGAAC
 4621 GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA
 4681 AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG
 4741 GGAGCTTCCA GGGGAAACG CCTGGTATCT TTATAGTCCT GTCGGGTTTC GCCACCTCTG
 4801 ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGCGG AGCCTATGGA AAAACGCCAG
 4861 CAACCGGCC TTTTACGGT TCCCTGGCCT TTGCTGGCCT TTGCTCACA TGTTCTTCC
 4921 TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTGAGTGAG CTGATACCGC
 4981 TCGCCGCAGC CGAACGACCG AGCGCAGCGA GTCAAGTGAGC GAGGAAGCGG AAGAGGCCCT
 5041 GATCGGGTAT TTTCTCCCTA CGCATCTGTG CGGTATTTC CACCGCATAT ATGGTGCAC
 5101 CTCAGTACAA TCTGCTCTGA TGCCGCATAG TTAAGCCAGT ATACACTCCG CTATCGCTAC
 5161 GTGACTGGGT CATGGCTGCG CCCCCGACACC CGCCAACACC CGCTGACGCC CCCTGACGGG
 5221 CCTGCTCTGCT CCCGGCATCC GCTTACAGAC AAGCTGTGAC CGTCTCCGGG AGCTGCATGT
 5281 GTCAGAGGT TTACCGTCA TCACCGAAC GCGCAGGGCA GCTGCGGTAA AGCTCATCAG
 5341 CGTGGTCGT AAGCGATTCA CAGATGTCTG CCTGTTCATC CGCGTCCAGC TCGTTGAGTT
 5401 TCTCCAGAAC CGTTAATGTC TGGCTCTGA TAAAGCGGGC CATGTTAAGG GCGGTTTTT
 5461 CCTGTTGGT CACTGATGCC TCCGTAAAG GGGGATTTCT GTTCAATGGG GTAATGATAC
 5521 CGATGAAACG AGAGAGGGATG CTCAGATAC GGGTTACTGA TGATGAACAT GCCCAGGTTAC
 5581 TGGAACTGTT TGAGGGTAAAC CAACTGGCGG TATGGATGCG GCGGGACCG AGAAAAAATCA
 5641 CTCAGGGTCA ATGCCAGGCC TTGCTTAATA CAGATGTAGG TGTCCACAG GGTAGCCAGC
 5701 AGCATCCCTGC GATGCAGATC CGGAACATAA TGGTGCAGGG CGCTGACTTC CGCGTTTCCA
 5761 GACTTACGA AACACGGAAA CGGAAGACCA TTCATGTTGT TGCTCAGGTC GCAGACGTTT
 5821 TGCAGCAGCA GTCGCTTCAC GTTCGCTCGC GTATCGGTGA TTCATTCTGC TAACCACTAA
 5881 GGCAACCCCG CCAGCCTAGC CGGGCTCTA ACGACAGGAG CACGATCATG CGCACCCGTG
 5941 GCCAGGACCC AACGCTGGCC GAGATGCAGC GCGTGCAGGGCT GCTGGAGATG GCGGACGCC
 6001 TGGATATGTT CTGCAAGGG TTGGTTTGC CATTACAGT TCTCCGCAAG AATTGATTGG
 6061 CTCCAATTCT TGGAGTGGTG AATCCGTTAG CGAGGTGCCG CGGGCTTCCA TTCAGGTCGA
 6121 GGTGGCCCGG CTCCATGCAC CGCGACGCAA CGCGGGGAGG CAGACAAGGT ATAGGGCGGC-

FIGURE 35c

90/240

6181 GCCTACAATC CATGCCAACCGTTCCATGT GCTCGCCGAG GCGGCATAAA TCGCCGTGAC
6241 GATCAGCGGT CCAGTGATCG AAGTTAGCT GGTAAGAGCC GCGAGCGATC CTTGAAGCTG
6301 TCCCTGATGG TCGTCATCTA CCTGCCTGGA CAGCATGGCC TGCAACGCGG GCATCCCAGT
6361 GCCGCCGGAA GCGAGAAGAA TCATAATGGG GAAGGCCATC CAGCCTCGCG TCGCGAACGC
6421 CAGCAAGACG TAGCCCCAGCG CGTCGGCCGC CATGCCGGCG ATAATGGCCT GCTTCTCGCC
6481 GAAACGTTTG GTGGCGGGAC CAGTGACGAA GGCTTGAGCG AGGGCGTGCA AGATTCCGA
6541 TACCGCAAGC GACAGGCCGA TCATCGTCGC GCTCCAGCGA AAGCGGTCT CGCCGAAAAT
6601 GACCCAGAGC GCTGCCGGCA CCTGTCTTAC GAGTTGCATG ATAAAAGAAGA CAGTCATAAG
6661 TGCGGCAGCG ATAGTCATGC CCCGCGCCCA CGGGAAGGAG CTGACTGGGT TGAAGGCTCT
6721 CAAGGGCATE GGTGATCGA CGCTCTCCCT TATCGGACTC CTGCATTAGG AAGCAGCCCA
6781 GTAGTAGGTT GAGGCCGTTG AGCACCGCCG CCGCAAGGAA TGGTGCATGC AAGGAGATGG
6841 CGCCCAACAG TCCCCCGGCC ACGGGGCCTG CCACCATACC CACGCCGAAA CAAGCGCTCA
6901 TGAGCCCGAA GTGGCGAGCC CGATCTTCCC CATCGGTGAT GTCGGCGATA TAGGCCAG
6961 CAACCGCACC TGTGGCGCCG GTGATGCCGG CCACGATGCG TCCGGCGTAG AGG

FIGURE 351)

91/240

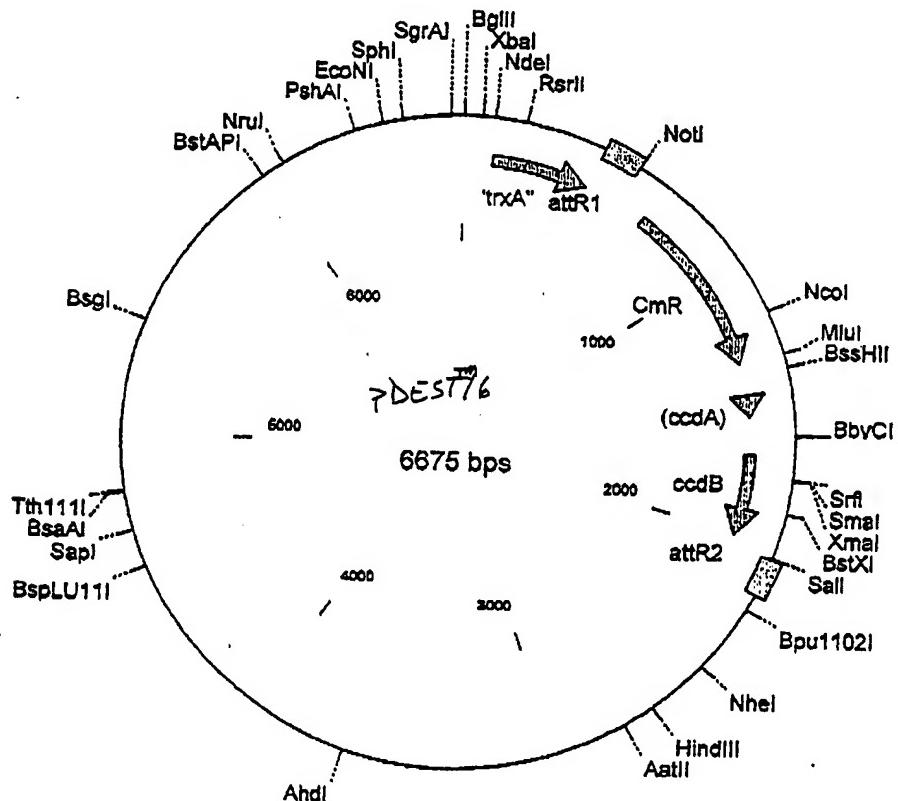
Figure 36A: *>DEST16***Thioredoxin N-Fusion Protein
in E. coli with T7 Promoter**

T7 Promoter mRNA →

```

1  gat ctc gat ccc gcg aaa tta ata cga ctc act ata [ggg] aga cca caa cgg
  cta gag cta ggg cgc ttt aat tat gct gag tga tat ccc tct ggt gtt gcc
  XbaI           NdeI
52   ttt ccc tct aga aat aat ttt gtt taa ctt taa gaa gga gat ata cat atc Start
  aaa ggg aga tct tta tta aaa caa att gaa att ctt cct cta tat gta gac Translation Trx
  S D K
103  agc gat aaa att att cac ctg act gac gac agt ttt gac acg gat gta ctc
  tcg cta ctt taa taa gtg gac tga ctg ctg tca aaa ctg tgc cta cat gag
  //---358
  gtg gcg gca acc aaa gtg ggt gca ctg tct aaa ggt cag ttg aaa gag ttc
  cac cgc cgt tgg ttt cac cca cgt gac aga ttt cca gtc aac ttt ctc aag
  T C G D D D K I
409  ctc gac gct aac ctg gcc ggt tct ggt tct ggt gat gac gat gac aag atc
  gag ctg cga ttg gac cgg cca aga cca aga cca cta ctg cta ctg ttc tag
  T S L Y K K A attR1
460  aca agt ttg tac aaa aaa gct gaa cga gaa acg taa aat gat ata aat atc
  tgt tca aac atg ttt ttt cga ctt gct ctt tgc att tta cta tat tta tag
  Int

```



92/240

pDEST16 6675 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
104..457	trxA
585..461	attR1
694..1353	CmR
1473..1557	inactivated ccdA
1695..2000	ccdB
2041..2165	attR2

1 AGATCTCGAT CCCCGAAAT TAATACGACT CACTATAGGG AGACCACAAAC GGTTTCCCTC
 61 TAGAAATAAT TTTGTAAAC TTTAAGAAGG AGATATACAT ATGAGCGATA AAATTATTCA
 121 CCTGACTGAC GACAGTTTG ACACGGATGT ACTCAAAGCG GACGGGGCGA TCCTCGTCGA
 181 TTTCTGGCA GAGTGGTGC GTCCTGCCA AATGATCGCC CGGATTCTGG ATGAAATCGC
 241 TGACGAATAT CAGGGCAAC TGACCGTTGC AAAACTGAAC ATCGATCAAA ACCCTGGCAC
 301 TGCGCCGAA TATGGCATCC GTGGTATCCC GACTCTGCTG CTGTTCAAAA ACGGTGAAGT
 361 GGCGGCAACC AAAGTGGGTG CACTGTCTAA AGGTCAGTTG AAAGAGTTCC TCGACGCTAA
 421 CCTGGCCGGT TCTGGTTCTG GTGATGACGA TGACAAGATC ACAAGTTTGT ACAAAAAAGC
 481 TGAACGAGAA ACGTAAAATG ATATAAAATAT CAATATATTA AATTAGATTT TGCATAAAAA
 541 ACAGACTACA TAATACTGTA AAACACAACA TATCCAGTC CTATGGCGGC CGCATTAGGC
 601 ACCCCAGGCT TTACACTTTA TGCTTCCGGC TCGTATAATG TGTTGATTTT GAGTTAGGAT
 661 CGGGCGAGAT TTTCAGGAGC TAAGGAAGCT AAAATGGAGA AAAAATCAC TGGATATACC
 721 ACCGTTGATA TATCCCAATG GCATCGAAA GAACATTTG AGGCATTTCA GTCAGTTGCT
 781 CAATGTACCT ATAACCAGAC CGTTCAGCTG GATATTACGG CCTTTTTAAA GACCGTAAAG
 841 AAAAATAAGC ACAAGTTTA TCCGGCCTT ATTACATTC TTGCCCCGCT GATGAATGCT
 901 CATCCGGAAT TCCGTATGGC AATGAAAGAC GGTGAGCTGG TGATATGGGA TAGTGTTCAC
 961 CCTTGTACCA CCGTTTCCA TGAGCAAAC GAAACGTTT CATCGCTCTG GAGTGAATAC
 1021 CACGACGATT TCCGGCAGTT TCTACACATA TATTGCAAG ATGTGGCGTG TTACGGTGA
 1081 AACCTGGCCT ATTTCCTAA AGGGTTTATT GAGAATATGT TTTCTGTCTC AGCCAATCCC
 1141 TGGGTGAGTT TCACCAAGTT TGATTTAAC GTGGCCAATA TGACAACTT CTTCGCCCCC
 1201 GTTTTCACCA TGGGCAAATA TTATACGCAA GGCACAAAGG TGCTGATGCC GCTGGCGATT
 1261 CAGGTTCATC ATGCCGTCTG TGATGGCTTC CATGTCGGCA GAATGCTTAA TGAATTACAA
 1321 CAGTACTGCG ATGAGTGGCA GGGGGGGCG TAAACGCGTG GATCCGGCTT ACTAAAAGCC
 1381 AGATAACAGT ATGCGTATT TGCGCCTGAT TTTTGGGTA TAAGAATATA TACTGATATG
 1441 TATACCCGAA GTATGTCAAA AAGAGGTGTG CTATGAAGCA GCGTATTACA GTGACAGTTG
 1501 ACAGCGACAG CTATCAGTTG CTCAAGGCAT ATATGATGTC AATATCTCCG GTCTGGTAAG
 1561 CACAACCATG CAGAATGAAG CCCGTCGTCT GCGTGGCAGA CGCTGGAAAG CGGAAAATCA
 1621 GGAAGGGATG GCTGAGGTGCG CCCGGTTTAT TGAAATGAAC GGCTTTTTG CTGACGAGAA
 1681 CAGGGACTGG TGAAATGCAG TTTAAGGTTT ACACCTATAA AAGAGAGAGC CGTTATCGTC
 1741 TGTTTGTGGA TGTACAGAGT GATATTATTG ACACGCCCGG GCGACGGATG GTGATCCCCC
 1801 TGGCCAGTGC ACGTCTGCTG TCAGATAAAG TCTCCCGTGA ACTTTACCCG GTGGTGCATA
 1861 TCGGGGATGA AAGCTGGCGC ATGATGACCA CCGATATGGC CAGTGTGCCG GTCTCCGTTA
 1921 TCGGGGAAGA AGTGGCTGAT CTCAGCCACC GCGAAATGA CATAAAAAAC GCCATTAACC
 1981 TGATGTTCTG GGGAAATAAA ATGTCAGGCT CCCTTATACA CAGCCAGTCT GCAGGTGAC
 2041 CATACTGACT GGATATGTT TGTTTACAG TATTATGTAG TCTGTTTTT ATGCAAATC
 2101 TAATTTAATA TATTGATATT TATATCATT TACGTTCTC GTTCAGCTT CTTGTACAAA
 2161 GTGGTGATGA TCCGGCTGCT AACAAAGCCC GAAAGGAAGC TGAGTTGGCT GCTGCCACCG
 2221 CTGAGCAATA ACTAGCATAA CCCCTGGGG CCTCTAAACG GGTCTTGAGG GTTTTTTGC
 2281 TGAAAGGAGG AACTATATCC GGATATCCAC AGGACGGGTG TGTCGCCAT GATCGCGTAG
 2341 TCGATAGTGG CTCCAAGTAG CGAAGCGAGC AGGACTGGC GGCGGCCAAA GCGGTCGGAC
 2401 AGTGCCTCGA GAACGGGTGC GCATAGAAAT TGCACTAACG CATATAGCGC TAGCAGCACG
 2461 CCATAGTGAC TGGCGATGCT GTCCGAATGG ACGATATCCC GCAAGAGGCC CGGCAGTACC
 2521 GGCATAACCA AGCCTATGCC TACAGCATCC AGGGTGACGG TGCCGAGGAT GACGATGAGC
 2581 GCATTGTTAG ATTTCATACA CGGTGCCGTA CTGCGTTAGC AATTTAACTG TGATAAAACTA
 2641 CCGCATTAAA GCTTATCGAT GATAAGCTGT CAAACATGAG AATTCTTGAA GACGAAAGGG
 2701 CCTCGTGATA CGCCTATTAA TATAGGTTAA TGTCATGATA ATAATGGTTT CTTAGACGTC
 2761 AGGTGGCACT TTTCGGGAA ATGTGCGCGG AACCCCTATT TGTTTATT TGCTAAATACA-

FIGURE 36B

93/240

2821 TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAT AATATTGAAA
 2881 AAGGAAGAGT ATGAGTATTG AACATTTCCG TGTGCCCTT ATTCCCTTT TTGCGGCATT
 2941 TTGCCTTCCT GTTTTTGCTC ACCCAGAAC GCTGGTAAAA GTAAAGATG CTGAAGATCA
 3001 GTTGGGTGCA CGAGTGGGTT ACATCGAAC GGATCTAAC AGCGGTAAGA TCCTTGAGAG
 3061 TTTTCGCCCC GAAGAACGTT TTCCAATGAT GAGCACTTT AAAGTTCTGC TATGTGGCGC
 3121 GGTATTATCC CGTGTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA
 3181 GAATGACTTG GTTGAGTACT CACCAAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT
 3241 AAGAGAATTA TGCAGTGCCTG CCATAACCAC GAGTGATAAC ACTGCGGCCA ACTTACTTCT
 3301 GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTG CACAACATGG GGGATCATGT
 3361 AACTCGCCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA
 3421 CACCAAGATG CCTGCAGCAA TGGCAACAAAC GTTGCAGCAA CTATTAACCTG GCGAACTACT
 3481 TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAAG TTGCAGGACC
 3541 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAACTCG GAGCCGGTGA
 3601 GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT GTAAAGCCCT CCCGTATCGT
 3661 AGTTATCTAC ACCACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA
 3721 GATAGGTGCC TCACGTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT
 3781 TTAGATGTAT TTAAAACCTTC ATTTTTAATT TAAAAGGATC TAGGTGAAGA TCCTTTTTGA
 3841 TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTT CACTGAGCGT CAGACCCCGT
 3901 AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTTCTG CGCGTAATCT GCTGCTTGCA
 3961 AACAAAAAAA CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT
 4021 TTTTCCGAAG GTAACCTGGCT TCAGCAGAGC GCAGATAACCA AATACTGTCC TTCTAGTGT
 4081 GCCGTAGTTA GGCCCACACT TCAAGAACTC TGTAGCACCG CCTACATACCC TCGCTCTGCT
 4141 AATCCTGTTA CCAGTGGCTG CTGCGAGTGG CGATAAGTCG TGTCTTACCG GGTTGGACTC
 4201 AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGCTGA ACGGGGGGTT CGTGCACACA
 4261 GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA
 4321 AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CGCGTAAGCG GCAGGGTCGG
 4381 AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCTGT
 4441 CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTTGTGA TGCTCGTCAG GGGGGCGGAG
 4501 CCTATGGAAA AACGCCAGCA ACAGCGCCCTT TTTACGGTTT CTGGCCTTTT GCTGGCCTTT
 4561 TGCTCACATG TTCTTTCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCCTT
 4621 TGAGTGAGCT GATAACGCTC GCGCGAGCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA
 4681 GGAAGCGGAA GAGCGCTGAC TGCGGTATTT TCTCCTTACG CATCTGTGCG GTATTTACA
 4741 CCGCATATAT GGTGCACTCT CAGTACAATC TGCTCTGATG CCGCATAGTT AAGCCAGTAT
 4801 ACACCTCCGCT ATCGCTACGT GACTGGTCA TGGCTGCGCC CCGACACCCCG CCAACACCCG
 4861 CTGACCGGCC CTGACGGGCT TGTCTGCTCC CGGCATCCGC TTACAGACAA GCTGTGACCG
 4921 TCTCCGGGAG CTGCAATGTT CAGAGGTTTT CACCGTCATC ACCGAAACGC GCGAGGCAGC
 4981 TGCGGTAAAG CTCATCAGCG TGGTCGTGAA GCGATTCACA GATGTCTGCC TGTTCATCCG
 5041 CGTCCAGCTC GTTGAGTTTC TCCAGAGCG TTAATGTCTG GCTTCTGTATA AAGCGGGCCA
 5101 TGTTAAGGGC GGTTTTTCTC TGTTGGTCA CTGATGCTC CCGTGTAAAGGG GGATTTCTGT
 5161 TCATGGGGGT AATGATAACCG ATGAAACCGAG AGAGGATGCT CACGATAACCG GTTACTGATG
 5221 ATGAACATGC CCGGTTACTG GAACGTTGTG AGGGTAAACA ACTGGCGGTAA TGGATGCCGC
 5281 GGGACCAGAG AAAAATCACT CAGGGTCAAT GCCAGCGCTT CGTTAATACA GATGTAGGTG
 5341 TTCCACAGGG TAGCCAGCAG CATCCCTGGCA TGCAGATCCG GAACATAATG GTGCAGGGCG
 5401 CTGACTTCCG CGTTTCCAGA CTTTACGAAA CACCGAAACCG GAAGGACCAATT CATGTTGTTG
 5461 CTCAGGTGCG AGACGTTTTG CAGCAGCAGT CGCTTCACGT TCGCTCGCGT ATCGGTGATT
 5521 CATTCTGCTA ACCAGTAAGG CAACCCCGCC AGCCTAGCCG GGTCTCAAC GACAGGAGCA
 5581 CGATCATGCG CACCCGTGGC CAGGACCCAA CGCTGCCGA GATGCGCCGC GTGCGGCCG
 5641 TGGAGATGGC GGACGCGATG GATATGTTCT GCCAAGGGTT GGTTTGCACCA TTACAGTTTC
 5701, TCCGCAAGAA TTGATTGGCT CCAATTCTTG GAGTGGTGAAC TCCGTTAGCG AGGTGCGGCC
 5761 GGCTTCCATT CAGGTCGAGG TGGCCCGCT CCATGCGACCG CGACGCAACG CGGGGAGGCA
 5821 GACAAGGTAT AGGGCGGCCGCTAATCCA TGCCAACCCG TTCCATGTGC TCGCCGAGGC
 5881 GGCATAAAATC GCCGTGACGA TCAGCGGTCC AGTGTACGAA GTTAGGCTGG TAAGAGCCGC
 5941 GAGCGATCCT TGAAGCTGTC CCTGATGGTC GTCATCTAC TGCCTGGACA GCATGGCCTG
 6001 CAACCGGGGC ATCCCCGATGC CGCCGGAAAGC GAGAAGAACG ATAATGGGA AGGCCATCCA
 6061 GCCTCCGCTC GCGAACGCCA GCAAGACGTA GCCCAGCGCG TCGCCGCCCA TGCCGGCGAT
 6121 AATGGCCTGC TTCTCGCCGA AACGTTGGT GGCAGGACCA GTGACGAAGG CTTGAGCGAG
 6181 GCGTGCAGG ATTCCGAATA CCGCAAGCGA CAGGGCGATC ATCGTCGCGC TCCAGCGAAA
 6241 GCGGTCTCG CCGAAAATGA CCCAGAGCGC TGCCGGCACC TGTCTACGA GTTGCATGAT-

FIGURE 36C

94/240

6301 AAAGAAGACA GTCATAAGTG CGCGGACGAT AGTCATGCC CGCGCCCACC GGAAGGAGCT
6361 GACTGGTTG AAGGCTCTCA AGGGCATCGG TCGATCGACG CTCTCCCTTA TGCGACTCCT
6421 GCATTAGGAA GCAGCCCAGT AGTAGGTTGA GGCGTTGAG CACCGCCGCC GCAAGGAATG
6481 GTGCATGCAA GGAGATGGCG CCCAACAGTC CCCCGGCCAC GGGGCCTGCC ACCATAACCA
6541 CGCCGAAACA AGCGCTCATG AGCCCCAAGT GGCGAGCCCG ATCTTCCCCA TCGGTGATGT
6601 CGGCATATA GGCGCCAGCA ACCGCACCTG TGGCGCCGGT GATGCCGGCC ACGATGCGTC
6661 CGGCGTAGAG GATCG

FIGURE 360

95/240

mRNA

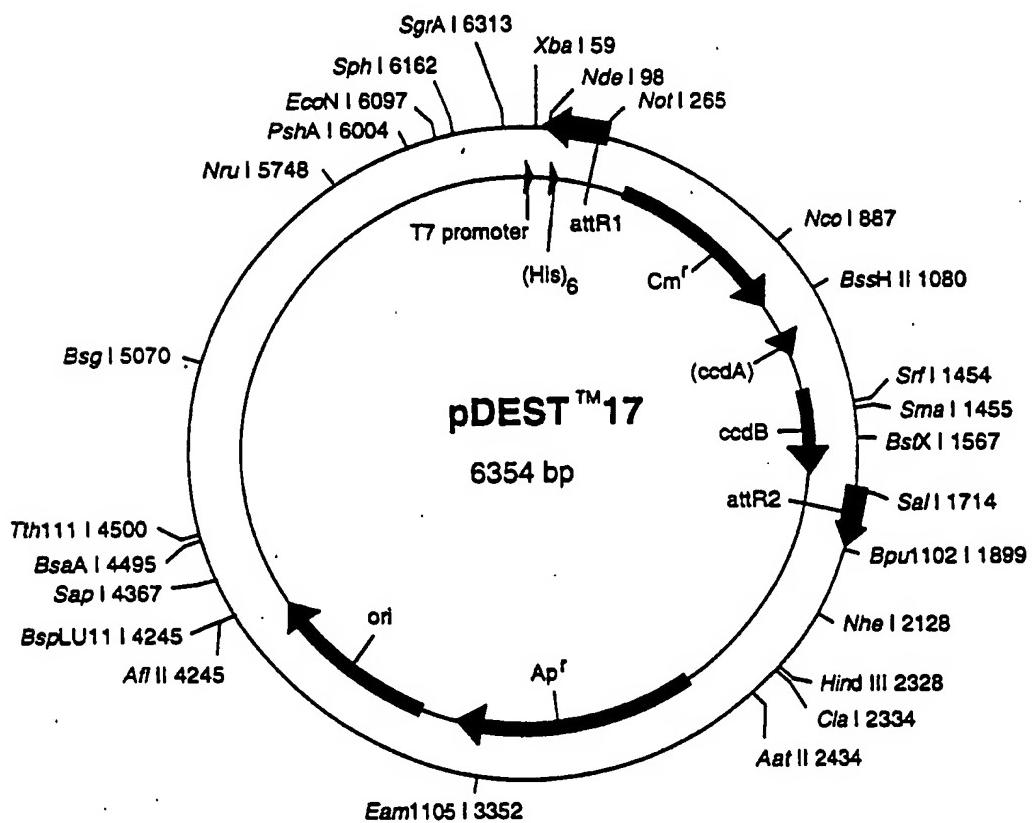
T7 Promoter

1 gat ccc gcg aaa tca ata cga ctc act ata ggg aga cca caa cgg ttt ccc
ctt ggg cgc ttt aat tat get gag tga tat ccc tct ggt gtt gcc aaa ggg

Start Translation M S Y

52 tct aga aat aat ttt gtt taa ctt taa gaa gga gat ata cat atg tgg tac
aga tct tta tta aaa caa att gaa att ctt cct cta tat gta tac aca atg
Y H H H H L E S T S L Y K K A //

103 tac cat cac cat cac cat cac ctc gaa tca aca agt tgg tac aaa aaa gct
atg gta gtg gta gtg gta gtg gag ctt agt tgg tca aac atg ttt ttt cga
attR1 Int V



96/240

pDEST17 6354 bp

Location (Base Nos.) Gene Encoded

258..134	attR1
367..1026	CmR
1146..1230	inactivated ccdA
1368..1673	ccdB
1714..1838	attR2
2564..3421	ampR

1 CGATCCCGCG AAATTAATAAC GACTCACTAT AGGGAGACCA CAACGGTTTC CCTCTAGAAA
 61 TAATTITGTT TAACTTAACG AAGGAGATAT ACATATGTCG TACTACCATC ACCATCACCA
 121 TCACCTCGAA TCAACAAGTT TGTACAAAAA AGCTGAACGA GAAACGTAAA ATGATATAAA
 181 TATCAATATA TTAAATTAGA TTTTGATCAA AAAACAGACT ACATAATACT GTAAAACACA
 241 ACATATCCAG TCACTATGGC GGCGCATTA GGCACCCAG GCTTTACACT TTATGCTTCC
 301 GGCTCGTATA ATGTGTGGAT TTTGAGTTAG GATCCGTCGA GATTTTCAGG AGCTAAGGAA
 361 GCTAAAAATGG AGAAAAAAAT CACTGGATAT ACCACCGTTG ATATATCCCA ATGGCATCGT
 421 AAAGAACATT TTGAGGCAATT TCAGTCAGTT GCTCAATGTA CCTATAACCA GACCGTTCA
 481 CTGGATATTAA CGGCCCTTTT AAAGACCGTA AAGAAAAATA AGCACAAGTT TTATCCGGCC
 541 TTTATTCAAC TTCTTGCCCCG CCTGATGAAT GCTCATCCGG AATTCCGTAT GGCAATGAAA
 601 GACGGTGAGC TGTTGATATG GGATAGTTGTT CACCTTGTT ACACCGTTTT CCATGAGCAA
 661 ACTGAAACGT TTTCATCGCT CTGGAGTGAA TACCAACGAG ATTTCGGCA GTTTCTACAC
 721 ATATATTTCGC AAGATGTGGC GTGTTACGGT GAAAACCTGG CCTATTTCCC TAAAGGGTTT
 781 ATTGAGAATA TGTTTTTCGT CTCAGCCAAT CCCTGGGTGA GTTTCACCAAG TTTTGATTTA
 841 AACGTGGCCA ATATGGACAA CTTCTTCGCC CCCGTTTCA CCATGGGCAA ATATTATACG
 901 CAAGGCAGCA AGGTGCTGAT GCGCCTGGCG ATTCAAGGTT ACATGCCGT CTGTGATGGC
 961 TTCCATGTCG GCAGAAATGCT TAATGAATTA CAACAGTACT GCGATGAGTG GCAGGGCGGG
 1021 GCGTAAAGAT CTGGATCCGG CTTACTAAAA GCCAGATAAC AGTATGCGTA TTTGCGCGCT
 1081 GATTTTTGCG GTATAAGAAT ATATACTGAT ATGTATACCC GAAGTATGTC AAAAAGAGGT
 1141 GTGCTATGAA GCAGCGTATT ACAGTGACAG TTGACAGCGA CAGCTATCAG TTGCTCAAGG
 1201 CATATATGAT GTCAATATCT CCGGTCTGGT AAGCACAACC ATGCAGAATG AAGCCCCTCG
 1261 TCTCGGTGCC GAACGCTGGA AAGCGGAAAA TCAGGAAGGG ATGGCTGAGG TCGCCCGGTT
 1321 TATTGAAATG AACGGCTCTT TTGCTGACGA GAACAGGGAC TGGTGAATG CAGTTTAAGG
 1381 TTTACACCTA TAAAAGAGAG AGCCGTTATC GTCTGTTTGT GGATGTACAG AGTGTATTTA
 1441 TTGACACGCC CGGGCGACGG ATGGTGTACCC CCCTGGCCAG TGCACGTCTG CTGTCAGATA
 1501 AAGTCTCCCCG TGAACCTTAC CCGGTGGTGC ATATGGGGGA TGAAAGCTGG CGCATGATGA
 1561 CCACCGATAT GGCCAGTGTG CCGGTCTCCG TTATGGGGGA AGAAGTGGCT GATCTCAGCC
 1621 ACCCGAAAAA TGACATCAA AACGCCATTA ACCTGATGTT CTGGGGAAATA TAAATGTCAG
 1681 GCTCCCTTAT ACACAGCCAG TCTGCAGGTC GACCATAGTG ACTGGATATG TTGTGTTTTA
 1741 CAGTATTATG TAGTCTGTTT TTATGCAAA ATCTAATTTA ATATATTGAT ATTTATATCA
 1801 TTTTACGTTT CTCGTTCAAC TTCTTGAC AAAGTGGTTG ATTGAGGCT GCTAACAAAG
 1861 CCCGAAAGGA AGCTGAGTTG GCTGCTGCCA CCGCTGAGCA ATAACCTAGCA TAACCCCTTG
 1921 GGGCCTCTAA ACGGGCTTTG AGGGGTTTT TGCTGAAAGG AGGAACATATA TCCGGATATC
 1981 CACAGGACGG GTGTGGTCGC CATGATCGCG TAGTCGATAG TGGCTCCAAG TAGCGAAGCG
 2041 AGCAGGACTG GGGGGCGGCC AAAGCGGTGCG GACAGTGCTC CGAGAACGGG TGCGCATAGA
 2101 AATTGCACTA ACGCATATAG CGCTAGCAGC ACGCCATAGT GACTGGCGAT GCTGTCGGAA
 2161 TGGACGATAT CCCCAGAGAG GCGCCGAGT ACCGGCATAA CCAAGCCTAT GCCTACAGCA
 2221 TCCAGGGTGA CGGTGCCAG GATGACGATG AGCGCATTGT TAGATTTCAT ACACGGTGCC
 2281 TGACTCGTGT AGCAATTAA CTGTGATAAA CTACCGCATT AAAGCTTATC GATGATAAGC
 2341 TGTCAAACAT GAGAATTCTT GAAGACGAAA GGGCCTCGTG ATACGCCATT TTTTATAGGT
 2401 TAATGTCATG ATAATAATGG TTCTTAGAC GTCAGGTGGC ACTTTTCGGG GAAATGTGCG
 2461 CGGAACCCCT ATTGTTTAT TTTCTAAAT ACATTCAAAT ATGTATCCGC TCATGAGACA
 2521 ATAACCCCTGA TAAATGCTTC AATAATATTG AAAAAGGAAG AGTATGAGTA TTCAACATTT
 2581 CCGTGTGCC CTTATTCCCT TTTTGCGGC ATTTTGCCCT CCTGTTTTG CTCACCCAGA
 2641 AACGCTGGTG AAAGTAAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA-

FIGURE 37B

97/240

2701 ACTGGATCTC AACAGCGGTA AGATCCTTGA GAGTTTCGC CCCGAAGAAC GTTTTCCAAT
 2761 GATGAGCACT TTTAAAGTTC TGCTATGTGG CGCGGTATT A TCCCGTGTG ACGCCGGCA
 2821 AGAGCAACTC GGTGCCGCA TACACTATTC TCAGAATGAC TTGGTTGAGT A CTCACCAGT
 2881 CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGCAGTG CTGCCATAAC
 2941 CATGAGTGAT AACACTGCGG CCAACTTAAT TCTGACAACG ATCGGAGGAC CGAAGGAGCT
 3001 AACCGCTTT TTGACAACA TGGGGGATCA TGTAACTCGC CTTGATCGTT GGGAAACCGGA
 3061 GCTGAATGAA GCCATACCAA ACAGACGAGCG TGACACCACG ATGCCTGCAG CAATGGCAAC
 3121 AACGTTGCGC AAACATTAA CTGGCGAATC ACTTACTCTA GCTTCCCAGC AACAAATTAA
 3181 AGACTGGATG GAGGGCGATA AAGTTGCAGG ACCACTCTG CGCTCGGCCC TTCCGGCTGG
 3241 CTGGTTTATT GCTGATAAT CTGGAGCGG TGAGCGTGGG TCTCGCGTA TCATTGCAGC
 3301 ACTGGGGCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG GGAGTCAGGC
 3361 AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA TTAAGCATTG
 3421 GTAATGTCA GACCAAGTTT ACTCATATAT ACTTAGATT GATTTAAAC TTCATTTTA
 3481 ATTTAAAAGG ATCTAGGTGA AGATCCTTT TGATAATCTC ATGACCAAAA TCCCTTAACG
 3541 TGAGTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA
 3601 TCCCTTTTTT CTGCGCGTAA TCTGCTGCTT GCAAACAAAA AAACCCACCGC TACCAAGCGG
 3661 GGTTTGTGAG CGGATCAAG AGCTACCAAC TCTTTTCCG AAGGTAACG GCTTCAGCAG
 3721 AGCGCAGATA CCAAATACTG TCCCTCTAGT GTAGCGCTAG TTAGGCCACC ACTTCAAGAA
 3781 CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAACTCTG TTACCAAGTGG CTGCTGCCAG
 3841 TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA TAGTTACCGG ATAAGGCGCA
 3901 GCGGTGGGGC TGAACGGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA CGACCTACAC
 3961 CGAACTGAGA TACCTACAGC GTGAGCTATG AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA
 4021 GGCAGCACAGG TATCCGGTAA GCGGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC
 4081 AGGGGGAAAC GCCTGGTATC TTATAGTCC TGTCGGGTTT CGCCACCTCT GACTTGAGCG
 4141 TCGATTTTTG TGATGCTCGT CAGGGGGGGC GAGCCTATGG AAAAACGCCA GCAACGCGG
 4201 CTTTTTACGG TTCCCTGGCCT TTTGCTGGCC TTTGCTCAC ATGTTCTTTC CTGCGTTATC
 4261 CCCTGATTCT GTGGATAACC GTATTACCGC CTTTGTAGTGA GCTGATAACCG CTCGCCCCAG
 4321 CCGAACGACC GAGCGCAGCG AGTCAGTGA CGAGGAAGCG GAAGAGCGCC TGATGGGTA
 4381 TTTCTCCTT ACCATCTGTG GCGGTATTC ACACCGATA TATGGTGCAC TCTCAGTACA
 4441 ATCTGCTCTG ATGCCGCATA GTTAAGCCAG TATACACTCC GCTATCGCTA CGTACTGGG
 4501 TCATGGCTGC GCCCGACAC CCGCCAACAC CCGCTGACGC GCCCTGACGG GCTTGTCTGC
 4561 TCCCGGCATC CGCTTACAGA CAAGCTGTGA CCGCTCCGG GAGCTGCATG TGTCAGAGGT
 4621 TTTCACCGTC ATCACCGAAA CGCGCGAGGC AGCTGCGGTAA AGCTCATCA GCGTGGCTGT
 4681 GAAGCGATTC ACAGATGTCT GCGTGTTCAT CCGCGTCCAG CTCGTTGAGT TTCTCCAGAA
 4741 GCGTTAATGT CTGGCTTCTG ATAAAGCGGG CCATGTTAAG GCGGGTTTTT TCCTGTTTGG
 4801 TCACTGATGC CTCCGTGTA GGGGGATTTC TGTTCATGGG GGTAAATGATA CCGATGAAAC
 4861 GAGAGAGGAT GCTCACGATA CGGGTTACTG ATGATGAACA TGCCCGGTTA CTGGAACGTT
 4921 GTGAGGGTAA ACAACTGGCG GTATGGATGC GCGGGGACCA GAGAAAAATC ACTCAGGGTC
 4981 AATGCCAGCG CTTGTTAAT ACAGATGTAG GTGTTCCACA GGGTAGCCAG CAGCATCTG
 5041 CGATGCAGAT CCGGAACATA ATGGTGCAGG GCGCTGACTT CCGCGTTTCC AGACTTTACG
 5101 AAACACGGAA ACCGAAGACC ATTCACTGTT TTGCTCAGGT CGCAGACGTT TTGCAGCAGC
 5161 AGTCGCTTC A CGTTCGCTCG CGTATCGGT ATTCACTCTG CTAACCAGTA AGGCAACCCC
 5221 GCCAGCCTAG CCGGGTCCTC AACGACAGGA GCACGATCAT GCGCACCCGT GGCCAGGACC
 5281 CAACGCTGCC CGAGATGCGC CGCGTGCAGC TGCTGGAGAT GCGCGACGCG ATGGATATGT
 5341 TCTGCCAAGG GTGGTTTGC GCAATTACAG TTCTCCGCAA GAATTGATTG GCTCCAATT
 5401 TTGGAGTGTG GAATCCGTTA GCGAGGTGCC GCGGGCTTCC ATTCAAGTGTG AGGTGGCCCG
 5461 GCTCCATGCA CCGCGACGCA ACAGGGGAG GCAGACAAGG TATAGGGCGG CGCCCTACAAT
 5521 CCATGCCAAC CCGTTCCATG TGCTCGCCGA GCGGGCATAA ATGCCCGTGA CGATCAGCGG
 5581 TCCAGTGTCA GAAAGTTAGGC TGGTAAGAGC CGCGAGCGAT CCTTGAAGCT GTCCCTGATG
 5641 GTCGTCATCT ACCTGCTCTG ACAGCATGGC CTGCAACCGC GGCATCCCGA TGCCGCGG
 5701 AGCGAGAAGA ATCATAATGG GGAAGGCCAT CCAGCCTCGC GTCGCGAACCG CCAGCAAGAC
 5761 GTAGCCCCAGC GCGTCGGCCG CCATGCCCCG GATAATGGCC TGCTTCTCGC CGAAACGTTT
 5821 GGTGGCGGGGA CCAAGTGCAGCA AGGCTTGAGC GAGGGCGTGC AAAGATTCCGA ATACCGCAAG
 5881 CGACAGGCCG ATCATCGTCG CGCTCCACCG AAAGCGGTCC TCGCCGAAAA TGACCCAGAG
 5941 CGCTGCCGGC ACCCTGCTCTA CGAGTTGCAT GATAAAAGAAG ACAGTCATAA GTGCGGGCAG
 6001 GATAGTCATG CCCCCGCGCCC ACCGGAAGGA GCTGACTGGG TTGAAGGCTC TCAAGGGCAT
 6061 CGGTCCATCG ACCGCTCTCCC TTATGCAGT CCGTGCATTAG GAAGCAGCCC AGTAGTAGGT
 6121 TGAGGCCGTT GAGCACCGCC GCGCAAGGA ATGGTGCATG CAAGGAGATG GCGCCCAACA-

FIGURE 37C

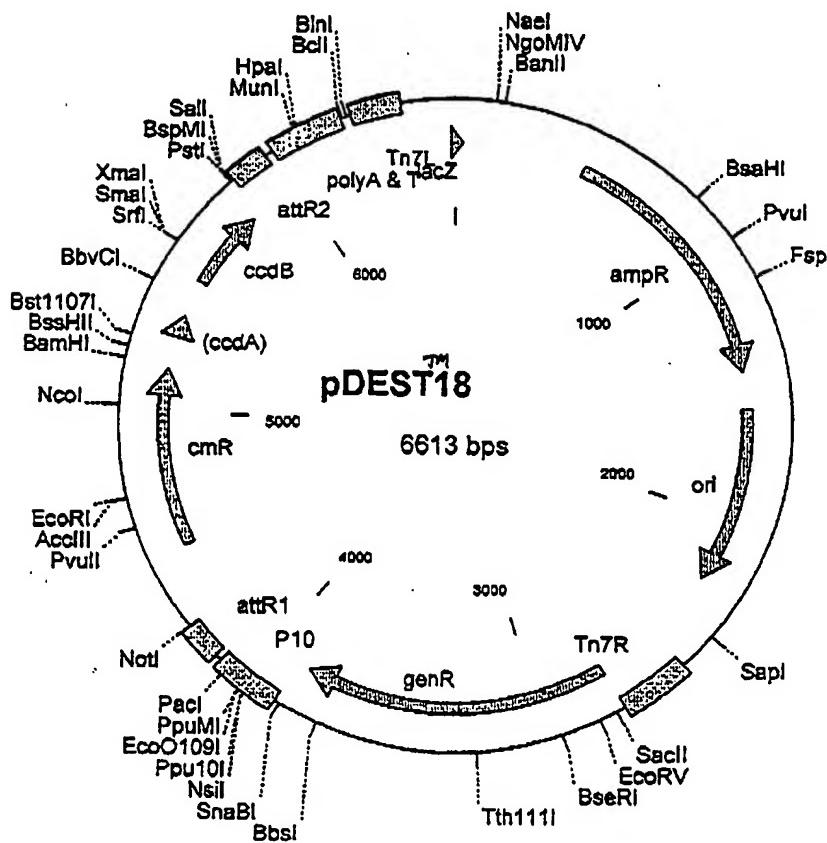
98/260

6181 GTCCCCCGGC CACGGGGCCT GCCACCATAAC CCACCCGAA ACAAGCGCTC ATGAGCCGA
6241 AGTGGCGAGC CCGATCTTCC CCATCGGTGA TGTGGCGAT ATAGGCGCCA GCAACCGCAC
6301 CTGTGGCGCC GGTGATGCCG GCCACGATGC GTCCCCCGTA GAGGATCGAG ATCT

FIGURE 37D

Figure 38A: DESTIB

FastBac Transfer Vector with p10 Baculovirus Promoter



100 / 260

pDEST18 6613 bp

Location (Base Nos.) Gene Encoded

474..1449	ampR
1590..2244	ori
2738..3850	genR
4251..4127	attR1
4501..5160	CmR
5280..5364	inactivated ccda
5502..5807	ccdB
5848..5972	attR2
6595..25	lacZ

1 GACGCGCCCT GTAGCGCGC ATTAAGCGCG GC GG GTGG TACGCG CAGCGTGACC
 61 GCTACACTTG CCAGCGCCCT AGCGCCCGT CCTTCGCCTT TCTTCCCTTC CTTCCTCGCC
 121 ACGTTGCCG GCTTCCCCG TCAAGCTCTA AATCGGGGGC TCCCTTAGG GTTCCGATTT
 181 AGTGTCTTAC GGCACCTCGA CCCCCAAAAAA CTTGATTAGG GTGATGGTTC ACGTAGTGGG
 241 CCATGCCCT GATAGACGGT TTTGCCCT TTGACGTTGG AGTCCACGTT CTTTAATAGT
 301 GGACTCTTGT TCCAAAATGG AACAAACACTC AACCCCTATCT CGGTCTATTTC TTTTGATTTA
 361 TAAGGGATTG TGCCGATTTG GGCCTATTGG TAAAAAAATG AGCTGATTAA ACAAAAAATTT
 421 AACGCGAATT TTAACAAAAT ATTAACGTTT ACAATTTCAG GTGGCACTTT TCAGGGAAAT
 481 GTGCGGGAA CCCCTATTG TTTATTTTC TAAATACATT CAAATATGTA TCCGCTCATG
 541 AGACAATAAAC CCTGATAAAAT GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA
 601 CATTCCGTG TCGCCCTTAT TCCCTTTTTG GCGGCATTG GCCTTCCGTG TTTTGCTCAC
 661 CCAGAAACGC TGGTGAAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTTAC
 721 ATCGAACTGG ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT
 781 CCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGGG TATTATCCC TATTGACGCC
 841 GGGCAAGAGC AACTCGGTG CCGCATAACAC TATTCTCAGA ATGACTTGTT TGAGTACTCA
 901 CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGTGCC
 961 ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG
 1021 GAGCTAACCG CTTTTTGCA CAACATGGGG GATCATGTA CTCGCTTGA TCGTTGGAA
 1081 CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG
 1141 GCAACAACGT TCGCAAACCT ATTAACTGGC GAACTACTTA CTCTAGCTTC CCGGCAACAA
 1201 TTAATAGACT GGATGGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG
 1261 GCTGGCTGGT TTATGCTGA TAAATCTGGA GCGGTGAGC GTGGGTCTCG CGGTATCATT
 1321 GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT
 1381 CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG
 1441 CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT AGATTGATT AAAACTTCAT
 1501 TTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTGATA ATCTCATGAC CAAAATCCCT
 1561 TAACGTGAGT TTTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTCT
 1621 TGAGATCCTT TTTTCTCGC CGTAATCTGC TGCTTGAAA CAAAAAAACC ACCGCTACCA
 1681 GCGGTGGTTT TTTGCCGGA TCAAGAGCTA CCAACTCTT TTCCGAAGGT AACTGGCTTC
 1741 AGCAGAGCGC AGATACAAA TACTGTCCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC
 1801 AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCCCTGTTACC AGTGGCTGCT
 1861 GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG
 1921 GCGCAGCGGT CGGGCTGAAC GGGGGGTTCG TGACACACAGC CCAGCTTGA GCGAACGACC
 1981 TACACCGAAC TGAGATACCT ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG
 2041 AGAAAGGCAG ACAGGTATCC GGTAAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG
 2101 CTTCCAGGGG GAAACGCCCTG GTATCTTAT AGTCTCTGTCG GGTTTCGCCA CCTCTGACTT
 2161 GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC
 2221 GCGGCCCTTT TACGGTTCCCT GGCCTTTGC TGGCCTTTTG CTCACATGTT CTTTCTCGCG
 2281 TTATCCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA TACCGCTCGC
 2341 CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCCTGATG
 2401 CGGTATTTTC TCCTTACGCA TCTGTGCGGT ATTTCACACC GCAGACCAGC CGCGTAACCT
 2461 GGAAACATCG GTTACGGTTG AGTAATAAT GGATGCCCTG CGTAAGCGGG TGTGGCGGA-

FIGURE 38B

2521 CAATAAAGTC TTAAACTGAA CAAAATAGAT CTAAAATATG ACAATAAAGT CTTAAAATAG
 2581 ACAGAAATAGT TGTTAAACTGA AATCAGTCGA GTTATGCTGT GAAAAAGCAT ACTGGAACTT
 2641 TGTTATGGCT AAAGCAAAC TCTCATTTC TGAAGTGCAA ATTGCCCCGTC GTATTAAAGA
 2701 GGGCGTGGC CAAGGGCATG GTAAAGACTA TATTGCGGGC GTTGTGACAA TTTACCGAAC
 2761 AACTCCGCGG CGGGGAAGCC GATCTCGGCT TGAACGAATT GTAGGTGGC GGTACTTGGG
 2821 TCGATATCAA AGTGCATCAC TTCTTCCCCT ATGCCAACT TTGTATAGAG AGCCACTGCG
 2881 GGATCGTCAC CGTAATCTGC TTGCACGTAG ATCACATAAG CACCAAGCGC STTGGCCTCA
 2941 TGCTTGAGGA GATTGATGAG CGCGGTGGC ATGCCCTGCC TCCGGTGTCT GCCGGAACT
 3001 GCGAGATCAT AGATATAGAT CTCACTACGC GGCTGCTCAA ACCTGGGAG AACGTAAGCC
 3061 GCGAGAGCGC CAACAACCGC TTCTTGGTGC AAGGCAGCAA GCGCGATGAA TGTCTTACTA
 3121 CGGAGCAAGT TCCCGAGGT ATCGGAGTCC GGCTGATGTT GGGAGTAGGT GGCTACGTCT
 3181 CCGAACTCAC GACCAGAAAG ATCAAGAGCA GCCCGCATGG ATTGACTTG GTCAGGGCCG
 3241 AGCCTACATG TGCGAATGTT GCCCATACTT GAGCCACCTA ACTTTGTTTT AGGGCGACTG
 3301 CCCTGCTGGC TAACATCGTT GCTGCTCGT AACATCGTT CTGCTCCATA ACATCAAACA
 3361 TCGACCCACG GCGTAACCGC TTGCTGCTT GGATGCCGA GCATAGACT GTACAAAAAA
 3421 ACAGTCATAA CAAGGCGATGA AAACCGCCAC TGCGCCGTTA CCACCGCTGC GTTCGGTCAA
 3481 GGTTCTGGAC CAGTTGCGTG AGCGCATACG CTACTTGCA TACAGTTTAC GAACCGAAACA
 3541 GGCTTATGTC AACTGGGTTG GTGCCCTTCAT CCGTTTCCAC GTGCGTGTCT ACCCGGCAAC
 3601 CTTGGGCCAG AGCGAAGTCG AGGCATTTCT GTCCCTGGCTG GCGAACCGAGC GCAAGGGTTT
 3661 GGTCTCCACG CATCGTCAGG CATTGGCGGC TTGCTGTTT TTCTACGGCA AGGTGCTGTG
 3721 CACGGATCTG CCCTGGCTTC AGGAGATCGG AAGACCTCGG CCCTGCGGGC GCTTGGCGGT
 3781 GGTGCTGACC CGGGATGAAG TGTTTCGAT CCTCGGTTTT CTGGAAGGGC AGCATCGTT
 3841 GTTGGCCAG GACTCTAGCT ATAGTTCTAG TGTTGGCTA CGTATCGAGC AAGAAAATAA
 3901 AACGCCAAC CGGTTGGAGT CTTGTGTGCT ATTGTTACAA AGATTCAAGA ATACGCATCA
 3961 CTTACAAACAA GGGGGACTAT GAAATTATGC ATTGAGGAA TGCCGGGAC TTTAATTCAA
 4021 CCCAACACAA TATATTATAG TTAAATAAGA ATTATTTATC AAATCATTG TATATTAAATT
 4081 AAAATACTAT ACTGTAATT ACATTTTATT TACAATGAGG ATCATCACAA GTTGTACAA
 4141 AAAAGCTGAA CGAGAAAAGT AAAATGATAT AAATATCAAT ATATTAATT AGATTTTGCA
 4201 TAAAAAACAG ACTACATAAT ACTGTAACAC ACAACATATC CAGTCACTAT GGCGGCCTGCT
 4261 AAGTTGGCAG CATCACCCGA CGCACTTGC GCCGAATAAA TACCTGTGAC GGAAGATCAC
 4321 TTCGCGAGAT AAATAAAATCC TTGTTGCTCT GTTGATACCG GGAAGCCCTG GGCAACCTTT
 4381 TGGCGAAAAT GAGACGTTGA TCGGCACGTA AGAGGTTCCA ACTTTCACCA TAATGAAATA
 4441 AGATCACTAC CGGGCGTATT TTGAGGTTA TCGAGATTTT CAGGAGCTAA GGAAGCTAAA
 4501 ATGGAGAAAA AAATCACTGG ATATACCACG TTGATATAT CCAATGGCA TCGTAAGAA
 4561 CATTGGAGG CATTTCAGTC AGTTGCTAA TGTACCTATA ACCAGACCGT TCAGCTGGAT
 4621 ATTACGGCCT TTGTTAAAGAC CGTAAAGAAA AATAAGCACA AGTTTATC GGCCTTATT
 4681 CACATTCTTG CCCGCCTGAT GAATGCTCAT CCGGAATTCC GTATGGCAAT GAAAGACGGT
 4741 GAGCTGGTGA TATGGGATAG TGTTCACCT TGTTACACCG TTGTTACATGA GCAAACATGAA
 4801 ACGTTTCAT CGCTCTGGAG TGAATACCAC GACGATTTC GGCAGTTCT ACACATATAT
 4861 TCGCAAGATG TGCGTGTGTTA CGGTGAAAAC CTGGCCTATT TCCCTAAAGG GTTGTGAG
 4921 AATATGTTTT TCGTCTCAGC CAATCCCTGG GTGAGTTCA CGAGTTTGA TTTAAACGTTG
 4981 GCCAATATGG ACAACATCTT CGCCCCCGTT TTCACCATGG GCAAATATTA TACGCAAGGC
 5041 GACAAGGTGC TGATGCCGCT GGCGATTCA GTCATCATG CCGTCTGTGA TGGCTTCCAT
 5101 GTCGGCAGAA TGCTTAATGA ATTACAAACAG TACTGCGATG AGTGGCAGGG CGGGGCSTAA
 5161 ACGCGTGGAT CGGGCTTACT AAAAGCCAGA TAACAGTATG CGTATTGCG CGCTGATT
 5221 TGCGGTATAA GAATATATAC TGATATGTAT ACCCGAAGTA TGTCAAAAAG AGGTGTCCTA
 5281 TGAAGCAGCG TATTACAGTG ACAGTTGACA GCGACAGCTA TCAGTTGCTC AAGGCAATA
 5341 TGATGTCAT ATCTCCGGTC TTGTTAAGCAG AACCATGCGAG ATGAAAGCCC GTCGCTGCG
 5401 TGCCGAAACGC TGAAAGCGG AAAATCAGGA AGGGATGGCT GAGGTGCGCC GGTGTTATTGA
 5461 AATGAACGGC TCTTTGCTG ACGAGAACAG GGACTGGTGA ATGCACTTT AAGGTTTACA
 5521 CCTATAAAAG AGAGAGCCGT TATCGTCGT TTGTTGGATGT ACAGAGTGAT ATTATTGACA
 5581 CGCCCCGGCG ACGGATGGTG ATCCCCCTGG CCAGTGCACG TCTGCTGTCA GATAAAGTCT
 5641 CCCGTAACCT TTACCCGGTG GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCAACCG
 5701 ATATGGCCAG TGTGCGGGTC TCCGTTATCG GGGAAAGAAGT GGCTGATCTC AGCCACCGCG
 5761 AAAATGACAT CAAAAACGCC ATTAACCTGA TGTTCTGGGG ATATATAATG TCAGGCTCCC
 5821 TTATACACAG CCAGTCTGCA GGTCGACCAT AGTGAATGGA TATGTTGTGT TTTACAGTAT
 5881 TATGTAATGCT GTTTTTATG CAAAATCTAA TTGTTATAT TGATATTTAT ATCATTTCAC
 5941 GTTTCGTT CAGCTTCTT GTACAAAGTG GTGATAGCTT GTCGAGAAGT ACTAGAGGAT-

FIGURE 38C

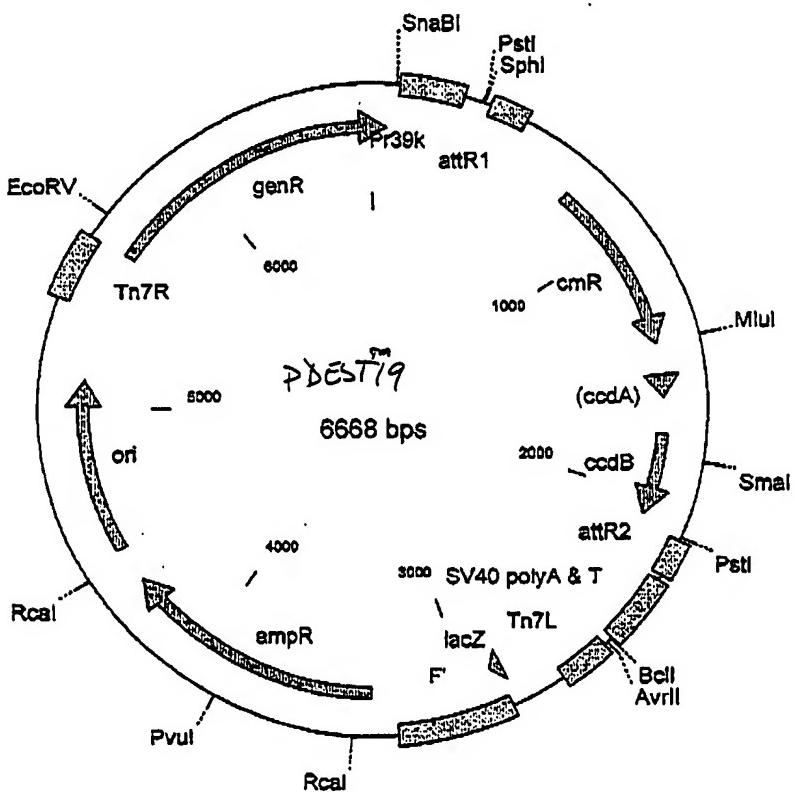
102/240

6001 CATAATCAGC CATACCACAT TTGTAGAGGT TTTACTTGCT TTAACCAACCT
6061 CCCCTGAAAC CTGAAACATA AAATGAATGC AATTGTTGTT GTAACTTGT TTATTGCAGC
6121 TTATAATGGT TACAATAAAA GCAATAGCAT CACAAATTTC ACAAAATAAG CATTTC
6181 ACTGCATTCT AGTTGTGGTT TGTCCAAACT CATCAATGTA TCTTATCATG TCTGGATCTG
6241 ATCACTGCTT GAGCCTAGGA GATCCGAACC AGATAAGTGA AATCTAGTTC CAAACTATTT
6301 TGTCACTTTT AATTTCGTA TTAGCTTAGC ACGCTACACC CAGTTCCCAT CTATTTGTC
6361 ACTCTTCCCT AAATAATCCT TAAAAACTCC ATTTCCACCC CTCCCAGTTC CCAACTATTT
6421 TGTCCGCCCA CAGCGGGGCA TTTTTCTTCC TGTTATGTTT TTAATCAAAC ATCCTGCCAA
6481 CTCCATGTGA CAAACCGTCA TCTTCGGCTA CTTTTCTCT GTACAGAAT GAAAATTTT
6541 CTGTCATCTC TTCGTTATTA ATGTTGTAA TTGACTGAAT ATCAACGCTT ATTTGCAGCC
6601 TGAATGGCGA ATG

FIGURE 38D

103/240

1 ggtgacggcc tcatcttccc attgtaacgt aaatggcaac ttgttagatga acgcgcgtgtc
 ccactcgccgc agtagaaaagg taacattgca ttatccgttg aacatctact tgcgcgacag
 61 aaaaaaacggg ccagtttccc ccacaaaactc gggcacgggt gtctcgtaaa cttttgcgtc
 tttttggcc ggtcaaagaa ggtgttttag cgctgcggca cagagcattt gaaaacgcag
 121 // gcaacaatcg cgatgacetc gggttatgga aaaaaaaaaaattttttctt aaaaaagtgt cgttcatgtc
 // cgttgttagc qctactqgag caccataacctt aaaaaaaga ttttttcaca gcaagtacag //
 181 // ggcggcgccggc ttgcgcgtcc ggtacggcg acgggcacac agcaggacag ctttgcgg
 // cccggccgc aagcgcgagg ccatgcgcgc tgccccgtgtc tcgtcccgatc ggaacaggcc
 241 ctcgattatc ataaaacaatc ctgcaggcat gcaagctgga tcatccaaag ttgtacaaa
 gagctaatacg tatttgttag gacgtccgtc cgttcgacact agtagtgcgtc aaacatgttt
 39.K Promoter
 39.R
 Int V



pDEST19 6668 bp (rotated to position 1000)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
515..391	attR1
765..1424	CmR
1544..1628	inactivated ccdA
1766..2071	ccdB
2112..2236	attR2
2852..2895	lacZ
3344..4319	ampR
4460..5114	ori
5608..52	genR

1 AGTGGTTCGC ATCCTCGGTT TTCTGGAAGG CGAGCATCGT TTGTTGCCCG AGGACTCTAG
 61 CTATAGTTCT AGTGGTTGGC TACGTATATC AAATACTTGT AGGTGACGCC GTCATCTTC
 121 CATTGTAACG TAAATGGCAA CTTGTAGATG AACGCGCTGT CAAAAAAACCG GCCAGTTCT
 181 TCCACAAACT CGCGCACGGC TGTCTCGTAA ACTTTGCGGT CGCAACAATC GCGATGACCT
 241 CGTGGTATGG AAATTTTTC TAAAAAAAGTG TCGTTCATGT CGCGGGCGGG CGCGTTCGCG
 301 CTCCGGTACG CGCGACGGGC ACACAGCAGG ACAGCCTTGT CCGGCTCGAT TATCATAAAC
 361 AATCTGCAAGC GCATGCAAGC TCGGATCATC ACAAGTTTGT ACAAAAAAGC TGAACGAGAA
 421 ACGTAAAATG ATATAAAATAT CAATATATTA AATTAGATTT TGATAAAAAA ACAGACTACA
 481 TAATACTGTA AAACACAACA TATCCAGTCA CTATGGCGGC CGCTAAGTTG GCAGCATCAC
 541 CCGACGCACT TTGCGCCGAA TAAATACCTG TGACGGAAGA TCACCTCGCA GAATAAATAA
 601 ATCCTGGTGT CCTGTGAT ACCGGGAAGC CCTGGGCCAA CTTTGGCGA AAATGAGACG
 661 TTGATCGGC CGTAAGAGGT TCCAACTTTC ACCATAATGA AATAAGATCA CTACCGGGCG
 721 TATTTTTGTA GTTATCGAGA TTTTCAGGAG CTAAGGAAGC TAAAATGGAG AAAAAAATCA
 781 CTGGATATAC CACCGTTGAT ATATCCCAAT GGATCGTAA AGAACATTTT GAGGCATTT
 841 AGTCAGTTGC TCAATGTACC TATAACCAGA CCGTTCACTG GGATATTACG GCCTTTTAA
 901 AGACCGTAAA GAAAAATAAG CACAAGTTT ATCCGGCTT TATTACATC CTTGCCGCC
 961 TGATGAATGC TCATCCGGAA TTCCGTATGG CAATGAAAGA CGGTGAGCTG GTGATATGGG
 1021 ATAGTGTCA CCCTGTTAC ACCGTTTCC ATGAGCAAAC TGAAACGTT TCATCGCTCT
 1081 GGAGTGAATA CCACGACGAT TTCCGGCAGT TTCTACACAT ATATTCGCAA GATGTGGCGT
 1141 GTTACGGTGA AAACCTGGCC TATTTCCCTA AAGGGTTTAT TGAGAATATG TTTTTCGTCT
 1201 CAGCCAATCC CTGGGTGAGT TTCACCACTT TTGATTTAAA CGTGGCCAAT ATGGACAAC
 1261 TCTTCGCCCC CGTTTCACC ATGGGCAAT ATTATACGCA AGCGGACAAG GTGCTGATGC
 1321 CGCTGGCGAT TCAGGTTCAT CATGCCGTCT GTGATGGCTT CCATGTCGGC AGAATGCTTA
 1381 ATGAATTACA ACAGTACTGC GATGAGTGGC AGGGCGGGGC STAAACGGCGT GGATCCGGCT
 1441 TACTAAAAGC CAGATAAACAG TATGCGTATT TGCGCGCTGA TTTTTCGGT ATAAGAATAT
 1501 ATACTGATAT GTATACCCGA AGTATGCAA AAAGAGGTGT GCTATGAAGC AGCGTATTAC
 1561 AGTGCACAGTT GACAGCGACA GCTATCAGTT GCTCAAGGCC TATATGATGT CAATATCTCC
 1621 GGTCTGGTAA GCACAACCAT GCAGAAATGAA GCCCGTCGTC TGCGTGGCGA ACGCTGGAAA
 1681 GCGGAAAATC AGGAAGGGAT GGCTGAGGTG GCCCGGTTTA TTGAAATGAA CGGCTCTTT
 1741 GCTGACGAGA ACAGGGACTG GTGAAATGCA GTTTAAGGTT TACACCTATA AAAGAGAGAG
 1801 CCGTTATCGT CTGTTTGTGG ATGTACAGAG TGATATTATT GACACGCCCG GGCACGGAT
 1861 GGTGATCCCC CTGGCCAGTG CACGTCGCT GTCAAGATAAA GTCTCCCGTG AACTTTACCC
 1921 GGTGGTGCAT ATCGGGGATG AAAGCTGGCG CATGATGACC ACCGATATGG CCAGTGTGCC
 1981 GGTCTCCGTT ATCGGGGAAAG AAGTGGCTGA TCTCAGCCAC CGGCAAATG ACATAAAAAA
 2041 CGCCATTAAC CTGATGTTCT GGGGAATATA AATGTCAGGC TCCCTTATAC ACAGCCAGTC
 2101 TGCAGGTGCA CCATAGTGAC TGGATATGTT GTGTTTACA GTATTATGTA GTCTGTTTT
 2161 TATGCAAAAT CTAATTTAAT ATATTGATAT TTATATCATT TTACGTTCT CGTTCAGCTT
 2221 TCTTGTACAA AGTGGTGCAT GAGAAGTACT AGAGGATCAT AATCAGCCAT ACCACATTG
 2281 TAGAGGTTT ACTTGCTTAA AAAAACCTCC CACACCTCCC CCTGAACCTG AAACATAAAA
 2341 TGAATGCAAT TGTTGTTGTT AACTTGTAA TTGCACTTA TAATGGTTAC AAATAAAGCA
 2401 ATAGCATCAC AAATTTCAAA AATAAAGCAT TTTTTCACT GCATTCTAGT TGTGGTTGT
 2461 CCAAACCTCAT CAATGTATCT TATCATGTC GGATCTGATC ACTGCTTGTAG CCTAGGAGAT
 2521 CCGAACCGA TAAGTGAAT CTAGTCCAA ACTATTTGT CATTAAAT TTTCGATTA
 2581 GCTTACGACG CTACACCCAG TTCCCATCTA TTTGTCACT CTTCCCTAAA TAATCCTAA-

FIGURE 39B

105/240

2641 AAACTCCATT TCCACCCCTC CCAGTTCCCA ACTATTTTGT CCGCCCACAG CGGGGCATTT
 2701 TTCTTCCTGT TATGTTTTTA ATCAAACATC CTGCCAACTC CATGTGACAA ACCGTCATCT
 2761 TCGGCTACTT TTTCTCTGTC ACAGAATGAA AATTTTTCTG TCATCTCTTC GTTATTAATG
 2821 TTTGTAATTG ACTGAATAATC AACGTTTATT TGCAGCCTGA ATGGCGAATG GACGCGCCCT
 2881 GTAGCGGCGC ATTAAGCGCG GCCTGGTGTGG TGTTACGCG CAGCGTGACC GCTACACTTG
 2941 CCAGCGCCCT AGCGCCCGCT CCTTTCCGTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCCCG
 3001 GCTTTCCCGC TCAAGCTCTA AATCGGGGGC TCCCTTTAGG GTTCCGATTIT AGTGTCTTAC
 3061 GGCACCTCGA CCCCCAAAAAA CTTGATTAGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT
 3121 GATAGACGGT TTTTCGCCCC TTGACGTTGG AGTCCACGTT CTTTAATAGT GGACTCTTGT
 3181 TCCAAAATCTGG AACAAACACTC AACCTATCT CCGGTTCTATTCT TTTTGATTTA TAAGGGATTT
 3241 TGCCGATTTG GGCCTATTGG TTAAAAAATG AGCTGATTTA ACAAAAATTT AACGCGAATT
 3301 TTAACAAAAT ATTAACGTTT ACAATTTCAG GTGGCACTTT TCAGGGAAAT GTGCGCGGAA
 3361 CCCCTATTG TTTATTCTTCAAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC
 3421 CCTGATAAAAT GCTTCAATAA TATTGAAAAA CGGAAGAGTAT GAGTATTCAA CATTTCGTC
 3481 TCGCCCTTAT TCCCTTTTT GCGGCATTTC GCCTTCTGT TTTTGCTCAC CCAGAACGC
 3541 TGGTGAAGT AAAAGATGCT GAAGATCACTGTTGGTGCACG AGTGGGTTAC ATCGAACTGG
 3601 ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTGCCCCCGA AGAACGTTT CCAATGATGA
 3661 GCACTTTAAAGTTCTGTA TGTGGCGCG TATTATCCCG TATTGACGCC GGGCAAGAGC
 3721 AACTCGGTG CCGCATAACAC TATTCTCAGA ATGACTTTGGT TGAGTACTCA CCAGTCACAG
 3781 AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGTGCGC ATAACCATGA
 3841 GTGATAAACAC TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG
 3901 CTTTTTGCA CAACATGGGG GATCATGAA CTCGCCCTGA TCGTTGGGAA CGGGAGCTGA
 3961 ATGAAGCCAT ACCAAACGAC GAGCGTACA CCACGATGCC TGTAGCAATG GCAACAAACGT
 4021 TGCACAAACT ATTAACGTTG GAACTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT
 4081 GGATGGAGGC GGATAAAAGTT GCAGGACAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT
 4141 TTATTGCTGA TAAATCTGGA GCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG
 4201 GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA
 4261 TGGATGAACG AAATAGACAG ATCGCTGAGA TAGTGCCTC ACTGATTAAG CATTGTTAAC
 4321 TGTCAGACCA AGTTTACTCA TATATACTTT AGATGATTT AAAACTTCAT TTTTAATTTA
 4381 AAAGGATCTA GGTGAAGATC CTTTTGATA ATCTCATGAC CAAATCCCT TAACGTGAGT
 4441 TTTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT
 4501 TTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA GCGGTGGTTT
 4561 TTTTGGCGGA TCAAGAGCTA CCAACTCTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC
 4621 AGATACCAAA TACTGTCCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG
 4681 TAGCACCGCC TACATACCTC GCTCTGCTAA TCCCTGTTACC AGTGGCTGCT GCCAGTGGCG
 4741 ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT
 4801 CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGGA GCGAACGACC TACACCGAAC
 4861 TGAGATACCT ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG
 4921 ACAGGTATCC GGTAAAGCGGC AGGGTGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG
 4981 GAAACGCCTG GTATCTTAT AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT
 5041 TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCCTTT
 5101 TACGGTTCCCT GGCCTTTTGC TGGCTTTTG CTCACATGTT CTTTCCTGCG TTATCCCCTG
 5161 ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAAGCTGA TACCGCTCGC CGCAGCGAA
 5221 CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCTGATG CGGTATTTTC
 5281 TCCTTACGCA TCTGTGCGGT ATTTCACACC GCAGACCAGC CGCGTAACCT GGCAAAATCG
 5341 GTTACGGTTG AGTAATAAAAT GGATGCCCTG CGTAAGCGGG TGTGGCGGA CAATAAAGTC
 5401 TTAAACTGAA CAAAATAGAT CTAAACTATG ACAATAAAGT CTAAACTAG ACAGAATAGT
 5461 TGTAAACTGAA ATCAGTCCA GTTATGCTGT GAAAAAAGCAT ACTGGACTTT TGTTATGGCT
 5521 AAAGCAAACT CTTCAATTTC TGAAGTGCCTA ATGCCCCGTC GTATTAAGA GGGGCGTGGC
 5581 CAAGGGCATG GTAAAGACTA TATTCGCGGC GTTGTGACAA TTTACCGAAC AACTCCGCGG
 5641 CCGGGAAAGCC GATCTCGGCT TGAACGAATT GTTGGTGGC GGTACTTGGG TCGATATCAA
 5701 AGTGCATCAC TTCTTCCCGT ATGCCAACT TTGTATAGAG AGCCACTGCG GGATCGTCAC
 5761 CGTAATCTGC TTGACGTAG ATCACATAAG CACCAAGCGC GTTGGCCTCA TGCTTGAGGA
 5821 GATTGATGAG CGCGGTGGCA ATGCCCTGCC TCCGGTGTCTC GCCGGAGACT GCGAGATCAT
 5881 AGATATAGAT CTCACTACGC GGCTGCTCAA ACCTGGGCAG AACGTAAGCC GCGAGAGCGC
 5941 CAACAACCGC TTCTTGGTCG AAGGCAGCAA GCGCGATGAA TGTCTTACTA CGGAGCAAGT
 6001 TCCCGAGGTA ATCGGAGTCC GGCTGATGTT GGGAGTAGGT GGCTACGTCT CCGAACTCAC
 6061 GACCGAAAAG ATCAAGAGCA GCCCGCATGG ATTTGACTTG GTCAGGGCCCG AGCCTACATG-

FIGURE 39C

106/240

6121 TGCAGATGAT GCCCATACTT GAGCCACCTA ACTTTGTTT AGGGCGACTG CCCTGCTGCG
6181 TAACATCGTT GCTGCTGCGT AACATCGTG CTGCTCCATA ACATCAAACA TCGACCCACG
6241 GCGTAACCGCG CTTGCTGCTT GGATGCCCGA GGCATAGACT GTACAAAAAA ACAGTCATAA
6301 CAAGCCATGA AAACCGCCAC TGCGCCGTTA CCACCGCTGC GTTCGGTCAA GGTTCTGGAC
6361 CAGTTGCGTG AGCGCATAACG CTACTTGCA TACAGTTAAC GAACCGAACA GGCTTATGTC
6421 AACTGGGTTC GTGCCCTTCAT CGGTTTCCAC GGTGTGCGTC ACCCGGCAAC CTTGGGCAGC
6481 AGCGAAGTCG AGGCATTCT GTCCTGGCTG GCGAACGAGC GCAAGGTTTC GGTCTCCACG
6541 CATCGTCAGG CATTGGCGGC CTTGCTGTT TTCTACGGCA AGGTGCTGTG CACGGATCTG
6601 CCCTGGCTTC AGGAGATCGG AAGACCTCGG CCGTCGCGGC GCTTGCCGGT GGTGCTGACC
6661 CCGGATGA

FIGURE 39D

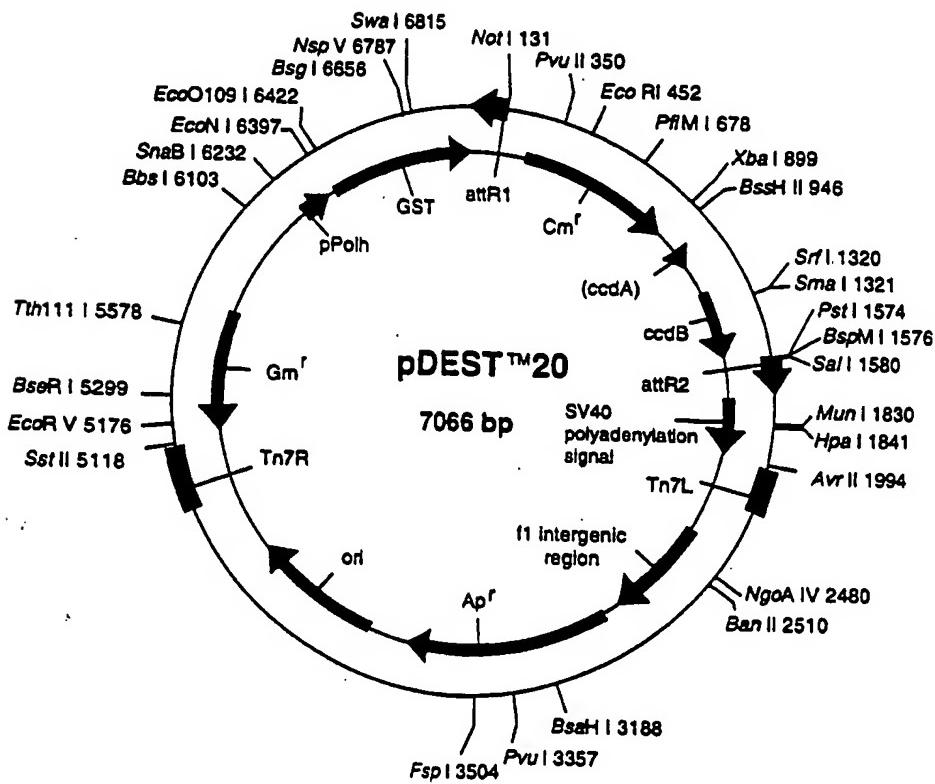
107/240

Figure 40A: pDEST20 Glutathione-S-transferase Fusion with Polyhedron Promoter for Baculovirus Expression

430 ggc tac gta tac tcc gga ata tta ata gat cat gga gat aat taa aat gat
 481 // aac cat ctc gca aat aaa taa gta ttt tac tgt ttt cgt aac aat ttt gta
 532 // atc aaa aaa cct ata aat att ccc gat tat tca tac cgt ccc acc atc ggg
 Start Transl. M → A P I GST -
 583 cgc gga tcc atg gct cct ata cta ggt tat tgg aaa att aag ggc ctt gtg
 gcg cct agg tac cgg gga tat gat cca ata acc ttt taa ttc cgg gaa cac

1246 S D L V P R H N Q T S L Y K R A
 // tcc gat ctg gtt ccc cgt cat dat caa aca agt tgg tac aaa aaa gct gaa
 agc cta gac caa ggc gca gta tta gtt tgt tca aac atg ttt ttt cga ctt

1297 cga gaa acg taa aat gat ata aat atc aat ata tta aat tag at
 gct ctt tgc att tta cta tat tta tag tta tat aat tta atc ta



108/240

pDEST20 7066 bp (rotated to position 5800)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
592..1263	GST
1397..1273	attR1
1506..2165	CmR
2285..2369	inactivated ccdA
2507..2812	ccdB
2853..2977	attR2
4214..5064	ampR
5263..5843	ori

1 CCACTGCGCC GTTACCAACCG CTGCCTTCGG TCAAGGTCT GGACCAGTTG CGTGAGCGCA
 61 TACGCTACTT GCATTACAGT TTACGAACCG AACAGGCTTA TGTCAACTGG GTTCGTCGCT
 121 TCATCCGTTT CCACGGTGTG CGTCACCCGG CAACCTTGGG CAGCAGCGAA GTCGAGGCAT
 181 TTCTGCTCTG GCTGGCGAAC GAGCGCAAGG TTTCGGTCTC CACGCATCGT CAGGCATTGG
 241 CGGCCTGCT GTTCTTCTAC GGCAAGGTGC TGTGACCGGA TCTGCCCTGG CTTCAGGAGA
 301 TCGGAAGACCC TCGGCCGTGCG CGGCCTTGC CGGTGGTGTG GACCCCGGAT GAAGTGGTTC
 361 GCATCCTCGG TTTCTGGAA GGCGAGCATC GTTGTTCGCG CCAGGACTCT AGCTATAAGTT
 421 CTAGTGGTTG GCTACGTATA CTCCGGAATA TTAATAGATC ATGGAGATAA TTAAAATGAT
 481 AACCATCTCG CAAATAATAA AGTATTTAC TGTTTCGTA ACAGTTTGT AATAAAAAAA
 541 CCTATAATAA TTCCGGATTAA TTCAACCGT CCCACCATCG GGCGCGGATC CATGGCCCC
 601 ATACTAGTT ATTGGAAAAT TAAGGGCCTT GTGCAACCCA CTCGACTTCT TTTGGAATAT
 661 CTTGAAGAAA AATATGAAGA GCATTTGTAT GAGCGCGATG AAAGGTGATAA ATGGCGAAAC
 721 AAAAGTTTG AATTGGGTTT GGAGTTTCCC AATCTTCCTT ATTATATTGA TGTTGATGTT
 781 AATAAACAC AGTCTATGGC CATCATACGT TATATAGCTG ACAAGCACAA CATGTTGGGT
 841 GGTTGTCAA AAGAGCGTGC AGAGATTCA ATGCTTGAAG GAGCGGTTTT GGATATTAGA
 901 TACGGTGTTC CGAGAAATTGC ATATAGTAA GACTTTGAAA CTCTCAAAGT TGATTTCTT
 961 AGCAAGCTAC CTGAAATGCT GAAAATGTT GAAGATCGT TATGTCATAA AACATATTAA
 1021 AATGGTGATC ATGTAACCCA TCCTGACTTC ATGTTGTATG ACGCTCTTGA TGTTGTTTA
 1081 TACATGGACC CAATGTGCCT GGATGCGTC CCAAAATTAG TTGTTTTAA AAAACGTATT
 1141 GAAGCTATCC CACAAATTGA TAAGTACTTG AAACTCCAGCA ACTATATAGC ATGGCCTTTG
 1201 CAGGGCTGGC AAGCCACGTT TGGTGGTGGC GACCATCCTC CAAAATCGGA TCTGGTTCCG
 1261 CGTCATAATC AAACAAGTTT GTACAAAAAA GCTGAACGAG AAACGTAAAAA TGATATAAAAT
 1321 ATCAATATAT TAAATTAGAT TTGCAATAA AAACAGACTA CATAATACTG TAAAACACAA
 1381 CATATCCAGT CACTATGGCG GCCGCATTAG GCACCCCAGG CTITACACTT TATGCTTCCG
 1441 GCTCGTATGT TGTGTGGATT TTGAGTTAGG ATCCGGCGAG ATTTCAGGA GCTAAGGAAG
 1501 CTAAAATGGA GAAAAAAATC ACTGGATATA CCACCGTTGA TATATCCAA TGGCATCGTA
 1561 AAGAACATTT TGAGGCATTT CAGTCAGTTG CTCAATGTAC CTATAACCAG ACCGTTTCAGC
 1621 TGGATATTAC GGGCTTTTTA AAGACCGTAA AGAAAAAATAA GCACAAGTTT TATCCGGCCT
 1681 TTATTACAT TCTTGGCCGC CTGATGAATG CTCATCCGGA ATTCCGTATG GCAATGAAAG
 1741 ACGGTGAGCT GGTGATATGG GATAGTGTTC ACCCTGTTA CACCGTTTTC CATGAGCAAA
 1801 CTGAAACGTT TTCATCGCTC TGGAGTGAAT ACCACGACGA TTTCCGGCAG TTTCTACACA
 1861 TATATTGCA AGATGTGGCG TGTACGGTG AAAACCTGGC CTATTTCCCT AAAGGGTTTA
 1921 TTGAGAAATAT GTTTTCGTC TCAGCCAATC CCTGGGTGAG TTTCACCAGT TTGATTAA
 1981 ACGTGGCCAA TATGGACAAC TTCTTCGCCC CCGTTTCAC CATGGCAAA TATTATAACGC
 2041 AAGGGACAA GGTGCTGATG CCGCTGGCGA TTCAAGGTTCA TCATGCCGT TGTGATGGCT
 2101 TCCATGTCGG CAGAATGCTT AATGAATTAC AACAGTACTG CGATGAGTGG CAGGGGGGG
 2161 CGTAATCTAG AGGATCCGGC TTACTAAAAG CCAGATAACA GTATGCGTAT TTGCGCGCTG
 2221 ATTTTGCAG TATAAGAATA TATACTGATA TGTATACCCG AAGTATGTCA AAAAGAGGTG
 2281 TGCTATGAAG CAGCGTATTA CAGTGACAGT TGACAGCGAC AGCTATCAGT TGCTCAAGGC
 2341 ATATATGATG TCAATATCTC CGGTCTGGTA AGCACAACCA TGCAGAAATGA AGCCCGTCGT
 2401 CTGCGTGGCG AACGCTGGAA AGCGGAAAT CAGGAAGGGG TGGCTGAGGT CGCCCGGTTT
 2461 ATTGAAATGA ACGGCTCTT TGCTGACCGAG AACAGGGACT GGTGAAATGC AGTTTAAGGT
 2521 TTACACCTAT AAAAGAGAGA GCGCTTATCG TCTGTTTGTG GATGTACAGA GTGATATTAT
 2581 TGACACGCC GGGCGACCGA TGGTGTACCC CCTGGCCAGT GCACGTCTGC TGTCAGATAA
 2641 AGTCTCCCGT GAACCTTACCG CGGTGGTGCA TATCGGGGAT GAAAGCTGGC GCATGATGAC-

Figure 40B

109/240

2701 CACCGATATG GCCAGTGTGC CGGTCTCCGT TATCGGGAA GAAGTGGCTG ATCTCAGCCA
 2761 CCGGAAAAT GACATAAAA ACGCCATTAA CCTGATGTT TGTTGGAAATAT AAATGTCAGG
 2821 CTCCCCTATA CACAGCCAGT CTGCAGGTG ACCATAGTGA CTGGATATGT TGTGTTTAC
 2881 AGTATTATGT AGTCTGTTTT TTATGCAAAA TCTAATTAA TATATTGATA TTTATATCAT
 2941 TTTACGTTTC TCGTTCAGCT TTCTTGACA AAGTGGTTTG ATAGCTTGTG GAGAAGTACT
 3001 AGAGGATCAT AATCAGCCAT ACCACATTG TAGAGGTTTG ACTTGCTTTA AAAAACCTCC
 3061 CACACCTCCC CCTGAACCTG AAACATAAAA TGAATGCAAT TGTTGTTGTT AACCTGTTA
 3121 TTGCAGCTTA TAATGGTTAC AAATAAAGCA ATAGCATCAC AAATTCACA AATAAAGCAT
 3181 TTTTTCACT GCATTCTAGT TGTGGTTTGT CCAAACCTCAT CAATGTATCT TATCATGTCT
 3241 GGATCTGATC ACTGCTTGAG CCTAGGAGAT CCGAACAGA TAAGTGAAT CTAGTCCAA
 3301 ACTATTTTGT CATTTTAAAT TTTCGTATTA GCTTACGACG CTACACCCAG TTCCCATCTA
 3361 TTTTGTCACT CTTCCCTAAA TAATCCTTAA AAACCTCATT TCCACCCCTC CCAGTTCCCA
 3421 ACTATTTTGT CCGCCCCACAG CGGGGCATTT TTCTTCTGT TATGTTTTA ATCAAACATC
 3481 CTGCAACTC CATGTGACAA ACCGTCACTC TCGGCTACTT TTTCTTGTG ACAGAATGAA
 3541 AATTTTTCTG TCATCTCTTC GTTATTAAAT TTTGTAATTG ACTGAATATC AACGCTTATT
 3601 TGCAGCCTGA ATGGCGAATG GACGCGCCCT GTAGCGGCCG ATTAAGCGCG GCGGGTGTGG
 3661 TGGTTACGCG CAGCGTGACC GCTACACTTG CCAGGCCCT AGCGCCCGT CCTTTGCTT
 3721 TCTTCCCTTC CTTTCTCGCC ACGTTGCCG GCTTCCCCG TCAAGCTCTA AATCGGGGGC
 3781 TCCCTTTAGG GTTCCGATTG AGTGCTTAC GGCACCTCGA CCCCCAAAAAA CTTGATTAGG
 3841 GTGATGGTTG ACGTAGTGGG CCATGCCCT GATAGACGGT TTTTCCCT TTGACGTTGG
 3901 AGTCACGTT CTTTAATAGT GGACTCTTGT TCCAAACTGG ACAACACACTC AACCCCTATCT
 3961 CGGCTTATTC TTTGATTAA TAAGGATTG TGCCGATTTC GGCTTATTGG TTTAAATATG
 4021 AGCTGATTTA ACAAAAATTT AACGCGAATT TTAACAAAAT ATTAAACGTTT ACAATTTCAG
 4081 GTGGCACTTT TCGGGGAAAT GTGCGGGG CCCCTATTG TTTTATTTC TAAATACATT
 4141 CAAATATGTA TCCGCTCATG AGACAATAAAC CCTGATAAAAT GCTTCAATAA TATTGAAAAA
 4201 GGAAGAGTAT GAGTATTCAA CATTCCGTG TCGCCCTTAT TCCCTTTTTT GCGGCATTTT
 4261 GCCTTCTGT TTTGCTCAC CCAGAACGC TGTTGAAAGT AAAAGATGCT GAAGATCAGT
 4321 TGGGTGACCG AGTGGGTTAC ATCGAACATGG ATCTCAACAG CGTAAAGATC CTTGAGAGATT
 4381 TTCGCCCGA AGAACGTTT CCAATGATGA GCACTTTAA AGTTCTGCTA TGTGGCGCGG
 4441 TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTG CCGCATAACAC TATTCTCAGA
 4501 ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGCC ATGACAGTAA
 4561 GAGAATTATG CAGTGTGCC ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTCTGA
 4621 CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTGCA CAACATGGGG GATCATGTAA
 4681 CTCGCCTTGA TCGTTGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA
 4741 CCACGATGCC TGTAGCAATG GCAACAACTG TGCGCAAAC ATTAACTGGC GAAACTACTTA
 4801 CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAAGTT GCAGGACCAC
 4861 TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATGCTGA TAAATCTGGA GCCGGTGAGC
 4921 GTGGGTCTCG CGGTATCATT GCAGCACTGG GGGCAGATGG TAAGCCCTCC CGTATCGTAG
 4981 TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA
 5041 TAGGTGCCTC ACTGATTAAG CATTGTAAC TGTCAAGACCA AGTTTACTCA TATATACTTT
 5101 AGATTGATTIT AAAACTTCAT TTTTAATTAA AAAGGATCTA GGTGAAGATC CTTTTTGATA
 5161 ATCTCATGAC CAAAATCCCT TAACGTGAGT TTCTGTTCCA CTGAGCGTCA GACCCCGTAG
 5221 AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTCTGCG CGTAATCTGC TGCTTGAAA
 5281 CAAAAAAACC ACCGCTACCA GCGGTGGTTT GTTGCCTGGA TCAAGAGCTA CCAACTCTTT
 5341 TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACCAA TACTGTCCTT CTAGTGTAGC
 5401 CGTAGTTAGG CCACCACTTC AAGAACCTG TAGCACCAGCC TACATACCTC GCTCTGCTAA
 5461 TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA
 5521 GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGTTCG TGCACACAGC
 5581 CCAGCTTGGA GCGAACGACC TACACCGAAC TGAGATACCT ACAGCGTGAG CATTGAGAAA
 5641 GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GTTAAGCGGC AGGGTCGGAA
 5701 CAGGAGAGCG CACGAGGGAG CTTCCAGGG GAAACGCGCTG GTATCTTAT AGTCTGTCG
 5761 GTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC
 5821 TATGGAAAAA CGCCAGCAAC GCGGCCTTT TACGGTTCCCT GGCCTTTGC TGGCCTTTG
 5881 CTCACATGTT CTTTCTGCG TTATCCCCTG ATTCTGTGGA TAACCGTATT ACCGCTTTG
 5941 AGTGAAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG
 6001 AAGCGGAAGA GCGCCTGATG CGGTATTTTC TCCTTACGCA TCTGTGCGGT ATTTACACACC
 6061 GCAGACCAGC CGCGTAACCT GGCAAAATCG GTTACGGTTG AGTAATAAAT GGATGCCCTG
 6121 CGTAAGCGGG TGTGGCGGA CAATAAGTC TAAACTGAA CAAAATAGAT CTAAACTATG-

FIGURE 40C

110/240

6181 ACAATAAAAGT CTTAAACTAG ACAGAATAGT TGTAACGTGA AATCAGTCCA GTTATGCTGT
6241 GAAAAAGCAT ACTGGACTTT TGTTATGGCT AAAGCAAAC TCTCATTTC TGAAAGTGC
6301 ATTGCCCGTC GTATTAAGA GGGGCGTGGC CAAGGGCATG GTAAAGACTA TATTCGCGGC
6361 GTTGTGACAA TTTACCGAAC AACTCCGCGG CGGGGAAGCC GATCTCGGCT TGAAACGAATT
6421 GTTAGGTGGC GGTACTTGGG TCGATATCAA AGTGCATCAC TTCTTCCCCT ATGCCCAACT
6481 TTGTATAGAG AGCCACTGCG GGATCGTCAC CGTAATCTGC TTGCACGTAG ATCACATAAG
6541 CACCAAGCGC GTTGGCCTCA TGCTTGAGGA GATTGATGAG CGCGGTGGCA ATGCCCTGCC
6601 TCCGGTGCTC GCCGGAGACT GCGAGATCAT AGATATAGAT CTCACTACGC GGCTGCTCAA
6661 ACCTGGGCAG AACGTAAGCC GCGAGAGCGC CAACAACCGC TTCTTGGTCC AAGGCAGCAA
6721 GCGCGATGAA TGTCTTAACCA CGGAGCAAGT TCCCCAGGTA ATCGGAGTCC GGCTGATGTT
6781 GGGAGTAGGT GGCTACGTCT CGGAACCTCAC GACGGAAAAG ATCAAGAGCA GCCCGCATGG
6841 ATTTGACTTG GTCAAGGGCCG AGCCTACATG TGCGAATGAT GCCCATACTT GAGCCACCTA
6901 ACTTTGTTTT AGGGCGACTG CCCTGCTGCG TAACATCGTT GCTGCTGCGT AACATCGTTG
6961 CTGCTCCATA ACATCAAAACA TCGACCCACG GCGTAACGCG CTTGCTGCTT GGATGCCGA
7021 GGCATAGACT GTACAAAAAA ACAGTCATAA CAAGCCATGA AAACCG

FIGURE 40D

111/240

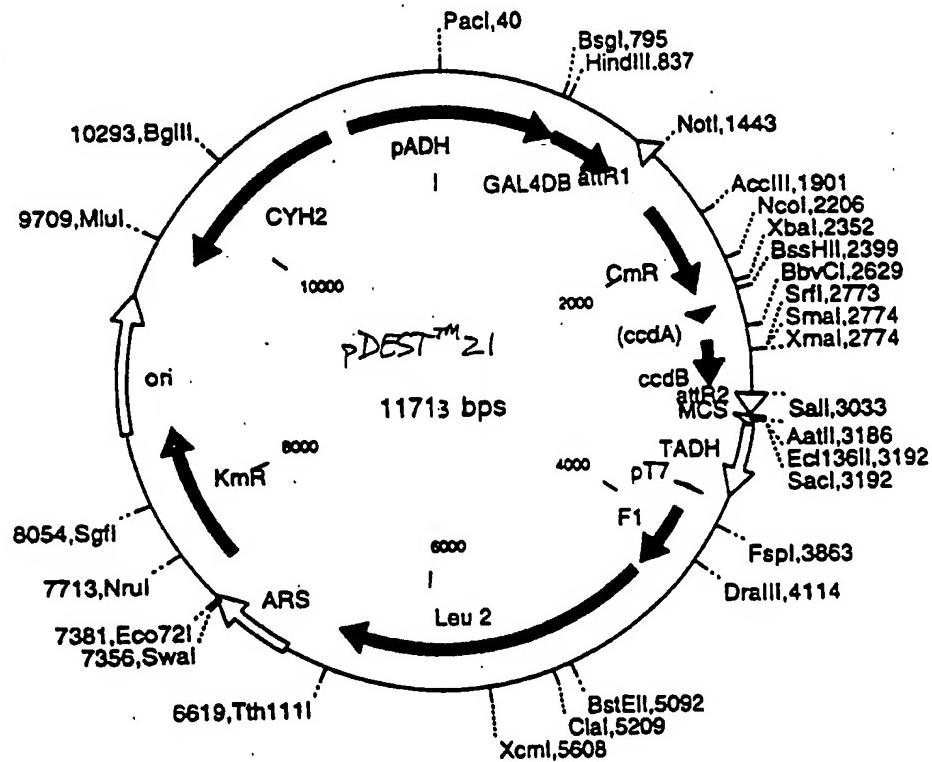
Figure 4(A):

pDEST21

2-Hybrid Vector with
DNA-Binding Domain

ADH Promoter

700 //ttg pcc ctc tcc taa gta taa atc gac ctg cca tta tta atc ttt tgt//
 751 //ttc ctc gtc att gtt ctc gtt ccc ttt cct tgc ttc ttc tgc aca//
 802 //ata ttt caa gct ata cca agc ata cca tca act cca aca ttg aag caa gcc//
 853 Start Transl M K L L S S Gal4 - DE
 tcc tga aag atg aag cca ctg tct tct atc gaa caa gca tgc gat att tgg//
 agg act ttc tac ttc gat gac aga tag ctt gtt cgt acg cta taa acg//
 ...
 1261 gaa gag agt agt aac aaa ggt caa aga cag ttg act gta tgg tgg agg tgg
 ctt ctc tca tca ttg ttt cca gtt tct gtc aac tga cat acg acg tcc acg
 1312 N Q T S L Y K K A R1
 dat caa aca agt tgg tac aaa aaa get gaa cga gaa acg taa aat gat ata
 tta gtt tgt tca aac atg ttt ttt cga ctt get ctt tgc att tta cta cat...//
 INT ↓



pDEST21 11713 bp (rotated to position 11000)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
857..1322	GAL4DB
1456..1332	attR1
1706..2365	CmR
2485..2569	inactivated ccdA
2707..3012	ccdB
3053..3177	attR2
3716..3735	pT7 (T7 promoter)
3899..4354	f1 (f1 intergenic region)
4414..6642	Leu2
7541..8515	kanR
9668..10958	CYH2
11118..848	pADH (ADH promoter)

1 TTTATTATGT TACAATATGG AAGGGAACCT TACACTTCTC CTATGCACAT ATATTAATTA
 61 AAGTCCAATG CTAGTAGAGA AGGGGGTAA CACCCCTCCG CGCTCTTTTC CGATTTTTT
 121 CTAAACCGTG GAATATTTCG GATATCCTTT TGTTGTTTCC GGGTGTACAA TATGGACTTC
 181 CTCTTTCTG GCAACCAAAC CCATACATCG GGATTCCTAT AATACCTTCG TTGGTCTCCC
 241 TAACATGTAG GTGGCGGAGG GGAGATATAC AATAGAACAG ATACCAGACA AGACATAATG
 301 GGCTAAACAA GACTACACCA ATTACACTGC CTCATTGATG GTGGTACATA ACGAACTAAT
 361 ACTGTAGCCC TAGACTTGTAG AGCCATCATC ATATCGAAGT TTCACTACCC TTTTCCATT
 421 TGCCATCTAT TGAAGTAATA ATAGGCGCAT GCAACTTCTT TTCTTTTTT TTCTTTCTC
 481 TCTCCCCCGT TGTGTCTCA CCATATCCGC AATGACAAAA AAAATGATGG AAGACACTAA
 541 AGGAAAAAAAT TAACGACAAA GACAGCACCA ACAGATGTG TGTTCCAGA GCTGATGAGG
 601 GGTATCTTCG AACACACGAA ACTTTTCTC TCCTTCATTC ACGCACACTA CTCTCTAATG
 661 AGCAACGGTA TACGGCCTTC CTTCCAGTTA CTTGAATTG AAATAAAAAAA AGTTTGCAGC
 721 TTTGCTATCA AGTATAAATA GACCTGCAAT TATTAATCTT TTGTTTCTC GTCATTGTT
 781 TCGTTCCCTT TCTTCCTTGT TTCTTTCTC GCACAATATT TCAAGCTATA CCAAGCATAC
 841 AATCAACTCC AAGCTTGAAG CAAGCCTCCT GAAAGATGAA GCTACTGTCT TCTATCGAAC
 901 AAGCATCGA TATTTGCCGA CTTAAAAAGC TCAAGTGCTC CAAAGAAAAA CCGAAGTGC
 961 CCAAGTGTCT GAAGAACAAAC TGGGAGTGTG GCTACTCTCC CAAAACCAAAG AGGTCTCCGC
 1021 TGACTAGGGC ACATCTGACA GAAGTGGAAAT CAAGGCTAGA AAGACTGGAA CAGCTATTC
 1081 TACTGATTTT TCCTCGAGAA GACCTTGACA TGATTTGAA AATGGATTCT TTACAGGATA
 1141 TAAAAGCATT GTTAAACAGGA TTATTTGTAC AAGATAATGT GAATAAAGAT GCGTCACAG
 1201 ATAGATTGGC TTCAGTGGAG ACTGATATGC CTCTAACATT GAGACAGCAT AGAATAAGTG
 1261 CGACATCATC ATCGGAAGAG AGTAGTAACA AAGGTCAAAG ACAGTTGACT GTATCGTC
 1321 GGTGAATCA ACAAGTTTG TACAAAAAAG CTGAACGAGA AACGTAACAT GATATAAATA
 1381 TCAATATATT AAATTAGATT TTGCATAAAA AACAGACTAC ATAATACTGT AAAACACAAAC
 1441 ATATCCAGTC ACTATGGCGG CCGCTAAGTT GGCAGCATCA CCCGACGCAC TTTGCGCCGA
 1501 ATAAAATACCT GTGACGGAAG ATCACTTCCG AGAATAAATA AATCCTGGTG TCCCTGTTGA
 1561 TACCGGGAAAG CCCTGGCCA ACTTTTGGCG AAAATGAGAC GTTGATCGGC ACGTAAGAGG
 1621 TTCCAACCTTT CACCATATG AAATAAGATC ACTACCGGGC GTATTTTTTG AGTTATCGAG
 1681 ATTTTCAGGA GCTAAGGAAG CTAAAATGGA GAAAAAAATC ACTGGATATA CCACCGTTGA
 1741 TATATCCCAA TGGCATCGTA AAGAACATT TGAGGCATT CAGTCAGTTG CTCAATGTAC
 1801 CTATAACCAAG ACCGTTCAAG TGGATAATTAC GGCTTTTTA AAGACCGTAA AGAAAAATAA
 1861 GCACAAGTTT TATCCGGCT TTATTCACAT TCTTGGCCGC CTGATGAATG CTCATCCGA
 1921 ATTCCGTATG GCAATGAAAG ACGGTGAGCT GGTGATATGG GATAGTGTTC ACCCTTGT
 1981 CACCGTTTC CATGAGCAAA CTGAAACGTT TTCATCGCTC TGAGGTGAAT ACCACGACGA
 2041 TTTCCGGCAG TTTCACACA TATATTCGCA AGATGTGGCG TGTACGGTG AAAACCTGGC
 2101 CTATTTCCCTT AAAGGGTTTA TTGAGAATAT GTTTTCTGTC TCAGCCAATC CCTGGGTGAG
 2161 TTTCACCAAGT TTTGATTAAAC ACCTGGCCAA TATGGACAAC TTCTTCGCCCC CCGTTTTCAC
 2221 CATGGGCAAA TATTATACGC AAGGCGACAA GGTGCTGATG CCGCTGGCGA TTCAGGTTCA
 2281 TCATGCCGTC TGTGATGGCT TCCATGTCCG CAGAACGCTT AATGAATTAC AACAGTACTG
 2341 CGATGAGTGG CAGGGCGGGG CGTAATCTAG AGGATCCGGC TTACTAAAAG CCAGATAACA
 2401 GTATGCGTAT TTGCGCGCTG ATTTTGCGG TATAAGAATA TATAACTGATA TGTATACCG-

FIGURE 4B

113/260

2461 AAGTATGTCA AAAAGAGGTG TGCTATGAAG CAGCGTATTA CAGTGACAGT TGACAGCGAC
 2521 AGCTATCAGT TGCTCAAGGC ATATATGATG TCAATATCTC CGGTCTGGTA AGCACAAACCA
 2581 TGCAGAATGA AGCCCGTCGT CTGCGTGGCC AACGCTGGAA AGCGGAAAAT CAGGAAGGGA
 2641 TGGCTGAGGT CGCCCCGGTTT ATTGAAATGA ACGGCTCTT TGCTGACGAG AACAGGGACT
 2701 GGTGAAATGC AGTTTAAGGT TTACACCTAT AAAAGAGAGA GCCGTTATCG TCTGTTGTG
 2761 GATGTACAGA GTGATATTAT TGACACGCC GGGCGACGGA TGGTGATCCC CCTGGCCAGT
 2821 GCACGTCTGC TGTCAGATAA AGTCTCCGT GAACTTTACC CGGTGGTGC TATCAGGGAT
 2881 GAAAGCTGGC GCATGATGAC CACCGATATG GCCAGTGTGC CGGTCTCCGT TATCAGGGAA
 2941 GAAGTGGCTG ATCTCAGCCA CCGCGAAAAT GACATCAAAA ACGCCATTAA CCTGATGTT
 3001 TGGGAAATAT AAATGTCAGG CTCCCTTATA CACAGCCAGT CTGCAGGTGC ACCATAGTGA
 3061 CTGGATATGT TGTTTTTAC AGTATTATGT AGTCTGTTT TTATGCAAAA TCTAATTAA
 3121 TATATGTA TTTATATCAT TTACGTTTC TCGTTCTAG TTCTTGTC AAGTGGTTTG
 3181 ATGGCCGCTA AGTAAGTAAG ACGTGAGCT CTAAGTAAGT AACGCCGGCC ACCGCGGTGG
 3241 AGCTTGGAC TTCTTCGCCA GAGGTTTGTG CAAGTCTCCA ATCAAGGTTG TCGGCTTGT
 3301 TACCTGCCA GAAAATTAACG AAAAGATGGA AAAGGGTCAA ATCGTTGGTA GATACTTGT
 3361 TGACACTCT AAATAAGCGA ATTTCTTAG ATTTATGATT TTATTATTA AATAAGTTAT
 3421 AAAAAAATA AGTGTATAACA AATTTTAAAG TGACTCTTAG GTTTTAAAC GAAAATTCTT
 3481 ATTCTTGAGT AACCTTCTTC TGTTAGTCAG GTTGTCTTCT CAGGTATAGC ATGAGGTCGC
 3541 TCTTATTGAC CACACCTCTA CGCGCATGCC GAGCAAATGC CTGCAAATCG CTCCCCATT
 3601 CACCCAATTG TAGATATGCT AACTCCAGCA ATGAGTTGAT GAATCTCGGT GTGTATT
 3661 TGTCCCTAGA GGACAATACC TGTTGTAATC GTTCTTCCAC ACGGATCCCA ATTGCCCTA
 3721 TAGTGAGTCG TATTACAATT CACTGGCCGT CGTTTACAA CGTCGTGACT GGGAAAACCC
 3781 TGGCGTTACC CAACTTAATC GCCTTGCAGC ACATCCCCCT TTGCGCCAGCT GGCGTAATAG
 3841 CGAAGAGGCC CGCACCGATC GCCCTTCCCA ACAGTTGCCG AGCCTGAATG GCGAATGGAC
 3901 GCGCCCTGTA GCGGCCGATT AAGCGCGGCG GGTGTGGTGG TTACGCGCAG CGTGACCGCT
 3961 ACACTGCCA GCGCCCTAGC GCGCCCTCCT TTGCGCTTCT TCCCTTCCTT TCTCGCCACG
 4021 TTGCGGGCT TTCCCGCTA AGCTCTAAAT CGGGGGCTCC CTTTAGGGTT CCGATTAGT
 4081 GCTTACGGC ACCTCGACCC CAAAAAACTT GATTAGGGTG ATGGTTACG TAGTGGGCCA
 4141 TCGCCCTGAT AGACGGTTT TCGCCCTTTC ACAGTTGGAGT CCACGTTCTT TAATAGTGG
 4201 CTCTTGTCC AACTGGAA AACACTCAAC CCTATCTCGG TCTATTCTTT TGATTATCAA
 4261 GGGATTTGCG CGATTTCGGC CTATTGTTA AAAATGAGC TGATTAAACA AAAATTAAAC
 4321 GCGAATTAAAC AAAAATATT AACGTTACA ATTTCTGAT GCGGTATTTT CTCCTTACGC
 4381 ATCTGTGCGG TATTCACAC CGCATATCGA CGCGTCGAGG AGAACCTCTA GTATATCCAC
 4441 ATACCTAATA TTATTGCTT ATTAAAAATG GAATCGGAAC AATTACATCA AAATCCACAT
 4501 TCTCTTCAAATC ATCAATTGTC CTGTTACTTCC TTGTTCATGT GTGTTCAAAA ACGTTATATT
 4561 TATAGGATAA TTATACTCTA TTTCTCAACA AGTAATTGGT TGTTTGGCCG AGCGGTCTAA
 4621 GGCGCCTGAT TCAAGAAATA TCTTGACCAGC AGTTAACTGT GGGAAACTCTC AGGTATCGTA
 4681 AGATGCAAGA GTTCAATCT CTTAGCAACC ATTATTTTT TCCTCAACAT AACGAGAAC
 4741 CACAGGGCG CTATCGCACA GAATCAAATT CGATGACTGG AAATTTTTG TTAATTTCAG
 4801 AGGTCGCTG ACGCATATAC CTTTTCAAC TGAAAAATTG GGAGAAAAAG GAAAGGTGAG
 4861 AGGCCGAAC CGGTTTTCA TATAGAATAG AGAACGCTTC ATGACTAAAT GCTTGCATCA
 4921 CAATACTTGA AGTTGACAAT ATTATTTAAG GACCTATTGT TTTTCCAAT AGGTGTTAG
 4981 CAATCGTCTT ACTTTCTAAC TTTTCTTAC TTTTACATT CAGCAATATA TATATATATT
 5041 TCAAGGATAT ACCATTCTAA TGTCTGCCCG TATGTCCTGC CCTAAGAAGA TCGTCGTTT
 5101 GCCAGGTGAC CACGTTGGTC AAGAAAATCAC AGCCGAAGCC ATTAAGGTTT TTAAAGCTAT
 5161 TTCTGATGTT CGTTCCAATG TCAAGTTCGA TTTGCAAAT CATTAAATTG GTGGTGTG
 5221 TATCGATGCT ACAGGTGTCC CACTTCCAGA TGAGGCGCTG GAAGCCTCCA AGAAGGTTGA
 5281 TGCGGTTTG TTAGGTGCTG TGGGTGGTCC TAAATGGGGT ACCGGTAGTG TTAGACCTGA
 5341 ACAAGGTTTA CTAAAAATCC GTAAAGAACT TCAATTGTC GCCAACITAA GACCATGTAA
 5401 CTTTGCATCC GACTCTCTT TAGACTTATC TCCAATCAAG CCACAATTG CTAAAGGTAC
 5461 TGACTTCGTT GTTGTCAAGAG AATTAGTGGG AGGTATTTCAC TTTGGTAAGA GAAAGGAAGA
 5521 CGATGGTGAT GGTGTGCGCTT GGGATAGTGA ACAATACACC GTTCCAGAAG TGCAAAGAAT
 5581 CACAAGAATG GCCGCTTTCA TGGCCCTACA ACATGAGCCA CCTTGCCTA TTTGGTCTT
 5641 GGATAAAAGCT ATGTTTTGG CCTCTTCAAG ATTATGGAGA AAAACTGTGG AGGAAACCAT
 5701 CAAGAACGAA TTCCCTACAT TGAAGGTTCA ACATCAATTG ATTGATTCTG CCGCCATGAT
 5761 CCTAGTTAAG ACCCAACCC ACCTAAATGG TATTATAATC ACCAGCAACA TGTTTGGTGA
 5821 TATCATCTCC GATGAAGCCT CCGTTATCCC AGGTTCCTTG GGTTTGTG CATCTGCGTC
 5881 CTTGGCCTCT TTGCCAGACA AGAACACCGC ATTGGTTTG TACGAACCAT GCCACGGTTC-

FIGURE 41C

114/240

5941 TGCTCCAGAT TTGCCAAAGA ATAAGGTTGA CCCTATGCC ACTATCTTGT CTGCTGCAAT
 6001 GATGTTGAAA TTGTCATTGA ACTTGCCTGA AGAAGGTAAAG GCCATTGAAG ATGCAGTTAA
 6061 AAAGGTTTG GATGCAGGTA TCAGAACTGG TGATTTAGGT GGTTCCAACA GTACCACCGA
 6121 AGTCGGTGTGAT GCTGTCGCCG AAGAAGTTAA GAAAATTCTT GCTTAAAAAG ATTCTCTTTT
 6181 TTTATGATAT TTGACATAA ACTTTATAAA TGAAATTCAAT ATAGAACG ACACGAAATT
 6241 ACAAAATGGA ATATGTTCAT AGGGTAGACG AAACATATA CGCAATCTAC ATACATTAT
 6301 CAAGAAGGAG AAAAAGGAGG ATAGTAAAGG AATACAGGTA AGCAAATTGA TACTAATGGC
 6361 TCAACGTGAT AAGGAAAAAG AATTGCACTT TAACATTAAT ATTGACAAGG AGGAGGGCAC
 6421 CACACAAAAA GTTAGGTGTA ACAGAAAATC ATGAAACTAC GATTCCATAAT TTGATATTGG
 6481 AGGATTTCT CTAAAAAAA AAAAATACAA CAAATAAAA ACACCTCAATG ACCTGACCAT
 6541 TTGATGGAGT TTAAGTCAAT ACCTTCTGTA ACCATTCCC ATAATGGTGA AAGTTCCCTC
 6601 AAGAATTTTA CTCTGTCAGA AACGGCCCTTA CGACGTAGTC GATATGGTGC ACTCTCAGTA
 6661 CAATCTGCTC TGATGCCGCA TAGTTAAGCC AGCCCCGACA CCCGCCAACA CCCGCTGACG
 6721 CGCCCTGACG GGCTTGTCTG CTCCCCGGCAT CCGCTTACAG ACAAGCTGTG ACCGTCCTCG
 6781 GGAGCTGCAT GTGTCAGAGG TTTTCACCGT CATCACCGAA ACGCGCGAGA CGAAACGGCC
 6841 TCGTGATACG CCTATTTTTA TAGGTTAATG TCATGATAAT AATGGTTCT TAGGACGGAT
 6901 CGCTTGCCTG TAACTTACAC GCGCCTCGTA TCTTTTAATG ATGGAATAAT TTGGGAATT
 6961 ACTCTGTGTT TATTTTATTT TATGTTTTGT ATTTGGATTT TAGAAAGTAA ATAAAGAAGG
 7021 TAGAAGAGTT ACGGAATGAA GAAAAAAA TAAACAAAGG TTTAAAAAAAT TTCAACAAAAA
 7081 AGCGTACTTT ACATATATAT TTATTAGACA AGAAAAGCAG ATTAATAGA TATACATTG
 7141 ATTAACGATA AGTAAAATGT AAAATCACAG GATTTCGTG TGTGGTCTTC TACACAGACA
 7201 AGATGAAACA ATTCGGCATT AATACCTGAG AGCAGGAAGA GCAAGATAAA AGGTAGTATT
 7261 TGTTGGCGAT CCCCTAGAG TCTTTACAT CTTCGAAAAA CAAAAACTAT TTTTTCTTAA
 7321 ATTTCTTTT TTACTTTCTA TTTTTAATTT ATATATTAT ATAAAAAAAT TTAAATTATA
 7381 ATTATTTTA TAGCACGTGA TGAAAAGGAC CCAGGTGGCA CTTTCCGGG AAATGTGCGC
 7441 GGAACCCCTA TTGTTTATT TTTCTAAATA CATTCAAATA TGTATCCGCT CATGAGACAA
 7501 TAACCCTGAT AAATGTTCA ATAATGTC GCTCTGGCCC GTGTCCTAAA ATCTCTGATG
 7561 TTACATTGCA CAAGATAAA ATATATCATC ATGAAACAATA AAACGTCTG CTTACATAAA
 7621 CAGTAATACA AGGGGTGTTA TGAGCCATAT TCAACGGAA ACGTCTTGCT GGAGGCCGCG
 7681 ATTAATTCC AACATGGATG CTGATTATA TGGGTATAAA TGGGCTCGCG ATAATGTCGG
 7741 GCAATCAGGT GCGACAATCT TCGATTGTA TGGGAAGCCC GATGCGCCAG AGTTGGTTCT
 7801 GAAACATGGC AAAGGTAGCG TTGCAATGA TGTTACAGAT GAGATGGTCA GACTAAACTG
 7861 GCTGACGGAA TTATGCTC TTCCGACCAT CAAGCATTTT ATCCGTACTC CTGATGATGC
 7921 ATGGTTACTC ACCACTGCGA TCCGCGGGAA AACAGCATTC CAGGTATTAG AAGAATATCC
 7981 TGATTCTAGGT GAAAATATTG TTGATCGCCT GGCAGTGTTC CTGCGCCGGT TGCATTGAT
 8041 TCCTGTTGT AATTGTCCTT TTAACAGCGA TCGCGTATTT CGTCTCGCTC AGGCGCAATC
 8101 ACGAATGAAAT AACGGTTTGG TTGATCGCAG TGATTTGAT GACGAGCGTA ATGGCTGGCC
 8161 TGTTGAAACAA GTCTGGAAAG AAATGCATAC GCTTTGCCA TTCTCACCGG ATTCACTCGT
 8221 CACTCATGGT GATTTCAC TTGATAACCT TATTGAC GAGGGGAAAT TAATAGGTTG
 8281 TATTGATGTT GGACGAGTCG GAATCGCAGA CCGATACCG GATCTTGCCA TCCTATGGAA
 8341 CTGCTCGGT GAGTTTCTC TTTCATTACA GAAAAGGCTT TTCAAAAT ATGGTATTGA
 8401 TAATCTGAT ATGAAATAAT TGCACTTCA TTGATGCTC GATGAGTTTT TCTAATCAGA
 8461 ATTGGTTAAT TGTTGTAAC ACTGGCAGAG CATTACGCTG ACTTGACGGG ACGGCGCATG
 8521 ACCAAAATCC CTTAACGTGA GTTTCGTTC CACTGAGCGT CAGACCCCGT AGAAAAGATC
 8581 AAAGGATCTT CTTGAGATCC TTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA
 8641 CCACCGCTAC CAGCGGTGGT TTGTTGCGG GATCAAGAGC TACCAACTCT TTTCCGAAG
 8701 GTAATGGCT TCAGCAGAGC GCAGATACCA AATACGTCC TTCTAGTGTG GCCGTAGTTA
 8761 GGCCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT AATCTGTTA
 8821 CCAGTGGCTG CTGCCAGTGG CGATAAGTC TGTCTTACCG GTGGGACTC AAGACGATAG
 8881 TTACCGGATA AGGCGCAGCG GTGGGCTGA ACGGGGGGTT CGTGCACACA GCCCAGCTG
 8941 GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCATTGAGA AAGCGCCACG
 9001 CTTCCGAAG GGAGAAAGGC GGACAGGTAT CGGTAAAGCG GCAGGGTCGG AACAGGAGAG
 9061 CGCACGGGG AGCTTCCAGG GGGGAACGCC TGTTATCTT ATAGTCCTGT CGGGTTTCGC
 9121 CACCTCTGAC TTGAGCGTCG ATTTTGATGA TGCTCGTCAG GGGGGCCGAG CCTATGGAAA
 9181 AACGCCAGCA ACGCGGCCCTT TTACGGTTC CTGGCTTTT GCTGGCTTT TGCTCACATG
 9241 TTCTTCTCTG CGTTATCCCC TGATTCTGT GATAACCGTA TTACCGCCTT TGAGTGANCT
 9301 GATACCGCTC GCCGCAAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA
 9361 GAGGCCCAA TACGCAAACC GCCTCTCCCC GCGCGTTGGC CGATTCAATTA ATGCAGCTGG-

FIGURE 4D

115/240

9421 CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA ACGCAATTAA TGTGAGTTAC
 9481 CTCACTCATT AGGCACCCCA GGCTTACAC TTTATGCTTC CGGCTCTAT GTTGTGTGGA
 9541 ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG ACCATGATTA CGCCAAGCTC
 9601 GGAATTAACC CTCACTAAG GGAACAAAAG CTGGTACCGA TCCCGAGCT TGCAAATTAA
 9661 AGCCTCGAG CGTCCCCAAA CCTTCTCAAG CAAGGTTTC AGTATAATGT TACATCGTAA
 9721 CACCGCTCTG TACAGAAAAA AAAGAAAAAT TTGAAATATA ATAACGTTC TTAATACTAA
 9781 CATAACTATA AAAAATAAAA TAGGGACCTA GACTTCAGGT TGTCTAATC CTTCCCTTTTC
 9841 GGTTAGAGCG GATGTGGGG GAGGGCGTGA ATGTAAGCGT GACATAACTA ATTACATGAT
 9901 ATCGACAAAG GAAAAGGGGC CTGTTACTC ACAGGCTTT TTCAGTAGG TAATTAAGTC
 9961 GTTTCTGTCT TTTCTCTCT TCAACCCACC AAAGGCCATC TTGGTACTTT TTTTTTTTTT
 10021 TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
 10081 TTTTTTTTTT TTTTTTTTTT TCATAGAAAAT AATACAGAAAG TAGATGTTGA ATTAGATTAA
 10141 ACTGAAGATA TATAATTATG TGAAAATAC ATAGAGCTTT TTGTTGATGC GCTTAAGCGA
 10201 TCAATTCAAC AACACCCACCA GCAGCTCTGA TTTTTCTTC AGCCAACITG GAGACGAATC
 10261 TAGCTTGAC GATAACTGGA ACATTTGGAA TTCTACCCCTT ACCCAAGATC TTACCGTAAC
 10321 CGGCTGCCAA AGTGTCAATA ACTGGAGCAG TTTCCCTTAGA AGCAGATTTC AAGTATTGGT
 10381 CTCTCTTGTG TTCTGGGATC AATGTCCACA ATTTGTCCAA GTTCAAGACT GGCTTCCAGA
 10441 AATGAGCTTG TTGTTGTGG AAGTATCTCA TACCAACCTT ACCGAAATAA CCTGGATGGT
 10501 ATTTATCCAT GTTAATTCTG TGGTGATGTT GACCACCGGC CATAACCTCTA CCACGGGGGT
 10561 GCTTTCTGTG CTTACCGATA CGACCTTTAC CGGCTGAGAC GTGACCTCTG TGCTTCTAG
 10621 TCTTAGTGAA TCTGGAAGGC ATTCTTGATT AGTGGATGA TTGTTCTGGG ATTTAATGCA
 10681 AAAATCACTT AAGAAGGAAA ATCAACGGAG AAAGCAAACG CCATCTTAAA TATACGGGAT
 10741 ACAGATGAAA GGGTTTGAAC CTATCTGGAA AATAGCATTA AACAAGCGAA AAACCTGCGAG
 10801 GAAAATTGTT TCGCTCTCTG CGGGCTATTAC ACAGCGCCAGA GGAAAATAGG AAAAATAACA
 10861 GGGCATTAGA AAAATAATT TGATTTGGT AATGTGTGGG TCCCTGGTGT AAGATGTTAC
 10921 ATTGGTTACA GTACTCTTGT TTTTGCTGTG TTTTCGATG AATCTCCAAA ATGGTTGTTA
 10981 GCACATGGAA GAGTCACCGA TGCTAAGTTA TCTCTATGTA AGCTACGTGG CGTACTTTT
 11041 GATGAAGCCG CACAAGAGAT ACAGGATTGG CAACTGCAA TAGAATCTGG GGATCCCCC
 11101 TCGAGATCCG GGATCGAAGA AATGATGGTA AATGAAATAG GAAATCAAGG AGCATGAAGG
 11161 CAAAAGACAA ATATAAGGGT CGAACGAAAA ATAAAGTGAA AAGTGTGAT ATGATGTATT
 11221 TGGCTTGCCT GCGCCAAAA AACGAGTTA CGCAATTGCA CAATCATGCT GACTCTGTGG
 11281 CGGACCCCGCG CTCTTGCCTGG CCCGGCGATA ACAGCTGGGCG TGAGGCTGTG CCCGGCGGAG
 11341 TTTTTGCGC CTGCAATTTC CAAGGTTTAC CCTGCGCTAA GGGGCGAGAT TGGAGAAGCA
 11401 ATAAGAATGC CGGTTGGGGT TGCGATGATG ACGACCACGA CAACTGGTGT CATTATTTAA
 11461 GTTGGCGAAA GAACCTGAGT GCATTTGCAA CATGAGTATA CTAGAAGAAT GAGCCAAGAC
 11521 TTGCGAGACG CGAGTTTGCC GGTGGTGCAG ACAATAGAGC GACCAGTACCC TTGAAGGTGA
 11581 GACGCGCATA ACCGCTAGAG TACTTTGAAG AGGAAACAGC AATAGGGTTG CTACCAAGTAT
 11641 AAATAGACAG GTACATACAA CACTGGAAAT GGTTGTCTGT TTGAGTACGC TTTCAATTCA
 11701 TTTGGGTGTG CAC

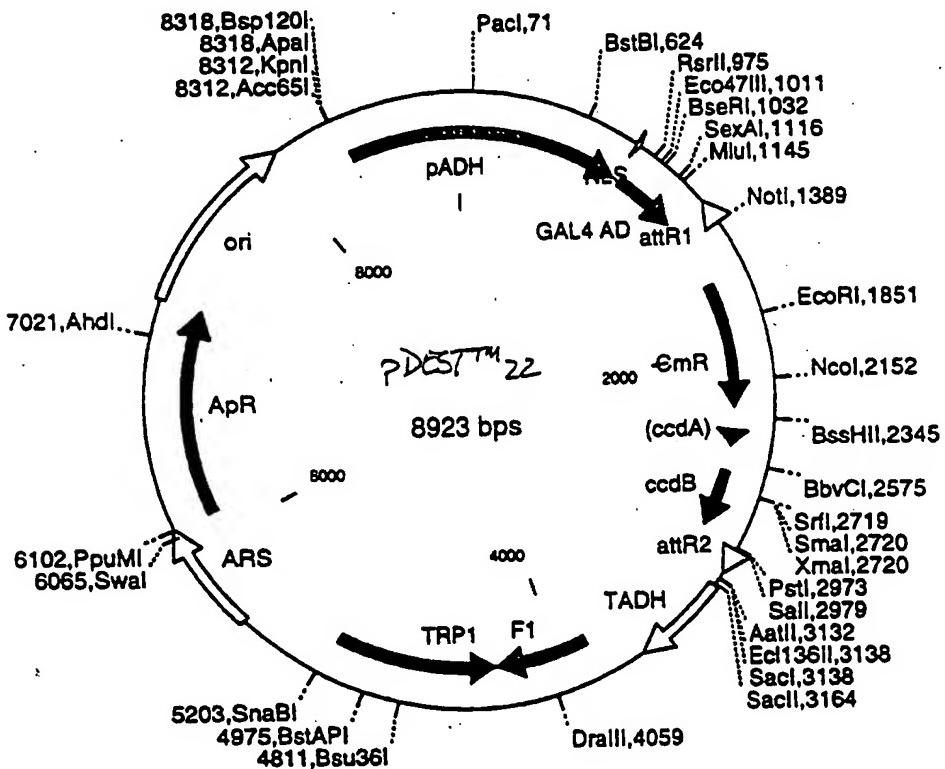
FIGURE 415

116/240

Figure 42A: PDEST22

2-Hybrid Vector with Activation Domain

657 acg cac act act ctc taa tga gca acg gta tac ggc ctt cct tcc agt tac
 tgc gtg tga tga gag att act cgt tgc cat atg ccg gaa gga agg tca atg
 708 ttg aat ttg aaa taa aaa aag ttt gec gct ttg cta tca agt ata aat aga
 aac tta aac ttt att ttt ttc aaa cgg cga aac gat agt tca tat tta tct
 759 cct gca att att aat ctt ttg ttt cct cgt cat tgt tct cgt tcc ctt tct
 gga cgt taa taa tta gaa aac aaa gga gca gta aca aga gca agg gaa aga
 810 // ~~ttc/ttg/ttt/ctt/ttc~~ ~~cgt~~ aat ~~att~~ ~~tcc~~ ~~epc~~ ~~tat~~ ~~acc~~ ~~aag~~ ~~cat~~ ~~acc~~ ~~atc~~
 // ~~aac~~ ~~aaa~~ ~~gaa~~ ~~aaa~~ ~~gac~~ ~~gtg~~ ~~tta~~ ~~aat~~ ~~agt~~ ~~tgg~~ ~~ata~~ ~~tgg~~ ~~tcc~~ ~~gtt~~ ~~tat~~ ~~tag~~
 861 // ~~aac~~ ~~ccc~~ ~~aag~~ ~~ctt~~ ~~atg~~ ~~ccc~~ ~~aag~~ ~~agg~~ ~~aag~~ ~~cg~~ ~~aag~~ ~~gtc~~ ~~tgg~~ ~~age~~ ~~ggc~~ ~~gcc~~ ~~aat~~
 // ~~tgg~~ ~~agg~~ ~~tcc~~ ~~gaa~~ ~~tac~~ ~~ggg~~ ~~tcc~~ ~~tcc~~ ~~gcc~~ ~~tcc~~ ~~cag~~ ~~age~~ ~~tgg~~ ~~ccg~~ ~~cg~~ ~~tta~~
 Start Translation
 1218 D G G S N Q T S
 gaa gat acc cca cca aac cca aaa aaa gag ggt ggt ggg tgg aat cca aca agt
 ctt cta tgg ggt ggt tgg ggt ttt ttt ctc cca ccc agc tta gtt tgt tca
 1269 // L Y K K A R I //
 // ~~tgg~~ ~~tac~~ ~~aaa~~ ~~aaa~~ ~~gct~~ ~~gaa~~ ~~cga~~ ~~gaa~~ ~~acg~~ ~~taa~~ a //
 // ~~aac~~ ~~atg~~ ~~ttt~~ ~~ttt~~ ~~cga~~ ~~ctt~~ ~~gct~~ ~~ctt~~ ~~tgc~~ ~~att~~ t //
 Infrv



117/260

pDEST22 8923 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
904..1248	GAL4 AD
1388..1264	attR1
1638..2297	CmR
2417..2501	inactivated ccdA
2639..2944	ccdB
2985..3109	attR2
3831..4318	f1 (f1 intergenic region)
4334..5176	TRP1
6110..7194	ampR
8344..866	pADH (yeast ADH promoter)

1 TTCATTGGG TGTGCACTTT ATTATGTTAC AATATGGAAG GGAACCTTAC ACTTCTCCTA
 61 TGACATATA TTAATTAAAG TCCAATGCTA GTAGAGAAGG GGGGTAACAC CCCTCCGC
 121 TCTTTCCGA TTTTTTCTA AACCGTGGAA TATTCGGAT ATCCTTTGT TGTTTCCGG
 181 TGTACAATAT GGACTTCCTC TTTTCTGGCA ACCAAACCCA TACATCGGGA TTCCCTATAAT
 241 ACCTTCGTTG GTCTCCCTAA CATGTAGGTG GCGGAGGGGA GATATACAAT AGAACAGATA
 301 CCAGACAGA CATAATGGGC TAAACAAGAC TACACCAATT ACACTGCCTC ATTGATGGTG
 361 GTACATAACG AACTAATACT GTAGCCCTAG ACTTGATAGC CATCATCATATA TCGAAGTTTC
 421 ACTACCCCTT TTCCATTGTC CATCTATTGA AGTAATAATA GGCGCATGCA ACTTCTTTTC
 481 TTTTTTTTC TTTTCTCTCT CCCCGTTGT TGTCTCACCA TATCCGCAAT GACAAAAAAA
 541 ATGATGGAAG ACACCAAAGG AAAAATTAA CGACAAAGAC AGCACCAACA GATGTCGTTG
 601 TTCCAGAGCT GATGAGGGGT ATCTTCGAAC ACACGAAACT TTTTCTTCC TTCATTACCG
 661 CACACTACTC TCTAATGAGC AACGGTATAC GGCCTCCCTT CCAGTTACTT GAATTTGAAA
 721 TAAAAAAAGT TTGCCGCTTT GCTATCAAGT ATAAATAGAC CTGCAATTAT TAATCTTTG
 781 TTTCCTCGTC ATTGTTCTCG TTCCCTTCT TCCTTGTTC TTTTCTGCA CAATATTCA
 841 AGCTATACCA AGCATAACAAT CAACTCCAAG CTTATGCCA AGAAGAAGCG GAAGGTCTCG
 901 AGCGGCAGCA ATTTTAATCA AAGTGGGAAT ATTGCTGATA GCTCATTGTC CTTCACTTTC
 961 ACTAACAGTA GCAACGGTCC GAACCTCATA ACAACTCAAA CAAATTCTCA AGCGCTTTCA
 1021 CAACCAATTG CCTCCTCTAA CGTTCATGAT AACTTCATGA ATAATGAAAT CACGGCTAGT
 1081 AAAATTGATG ATGGTAATAA TTCAAAACCA CTGTCACCTG GTTGGACCGA CCAAACGTGCG
 1141 TATAACCGT TTGGAATCAC TACAGGGATG TTTAATACCA CTACAATGGG TGATGTATAT
 1201 AACTATCTAT TCGATGATGA AGATAACCCA CAAACCCAA AAAAGAGGG TGGGTCGAAT
 1261 CAAACAAAGTT TGTACAAAAA AGCTGAACGA GAAACGTAAA ATGATATAAA TATCAATATA
 1321 TTAAATTAGA TTTTGCTAA AAAACAGACT ACATAATACT GTAAACACCA ACATATCCAG
 1381 TCACTATGGC GGCGCTAAG TTGGCAGCAT CACCGACGC ACTTTGCGGCC GAATAAAATAC
 1441 CTGTGACGGA AGATCACTTC GCAGAAATAA TAAATCTGG TGCCCTGTT GATACGGGA
 1501 AGCCCTGGC CAACTTTGG CGAAAATGAG ACCTGATCG GCACGTAAGA GGTTCCAAGT
 1561 TTCACCATAA TGAAAATAAGA TCACTACCCG GCGTATTTTG TGAGTTATCG AGATTTTCAG
 1621 GAGCTAAGGA AGCTAAAATG GAGAAAAAAA TCACTGGATA TACCACCGTT GATATATCCC
 1681 AATGGCATCG TAAAGAACAT TTGAGGCAT TTCACTCAGT TGCTCAATGT ACCTATAACC
 1741 AGACCGTTCA GCTGGATATT ACGGCCTTT TAAAGACCGT AAAGAAAAAT AAGCACAAGT
 1801 TTTATCCGGC CTTTATTCA ATTCTTGCCC GCCTGATGAA TGCTCATCCG GAATTCCTGTA
 1861 TGGCAATGAA AGACGGTGAG CTGGTGATAT GGGATAGTGT TCACCCCTGT TACACGGTTT
 1921 TCCATGAGCA AACTGAAACG TTTTCATCGC TCTGGAGTGA ATACCAACGAC GATTTCCGGC
 1981 AGTTTCTACA CATAATTCG CAAGATGTT CGTGTACCGG TGAACACCTG GCCTATTTC
 2041 CTAAAGGGTT TATTGAGAAT ATGTTTTCTG TCTCAGCCAA TCCCTGGGTG AGTTTCACCA
 2101 GTTTGATTT AAACGTGGCC AATATGGACA ACTTCTTCGC CCCCCTTTC ACCATGGGCA
 2161 AATATTATAC GCAAGCGAC AAGGTGCTGA TGCCGCTGGC GATTCAAGGTT CATCATGCC
 2221 TCTGTGATGG CTTCCATGTC GGCAGAAATGCTA TTAATGAATT ACAACAGTAC TGCGATGAGT
 2281 GGCAGGGCGG GGCAGTAATCT AGAGGATCCG GCTTACTAAA AGCCAGATAA CAGTATCGT
 2341 ATTTGCGCGC TGATTTTGC GGTATAAGAA TATATACTGA TATGTATACC CGAAGTATGT
 2401 CAAAAAGAGG TGTGCTATGA AGCAGCGTAC TACAGTGACA GTTGACAGCG ACAGCTATCA
 2461 GTTGCTCAAG GCATATATGA TGTCAATATC TCCGGTCTGG TAAGCACAAC CATGCAGAAT
 2521 GAAGCCGTC GTCTGCGTGC CGAACGCTGG AAAGCGGAAA ATCAGGAAGG GATGGCTGAG-

FIGURE 425

118/260

2581 GTCGCCCGGT TTATTGAAAT GAACGGCTCT TTTGCTGACG AGAACAGGG A CTGGTGAAAT
 2641 GCAGTTAAG GTTTACACCT ATAAAAGAGA GAGCCGTTAT CCTCTGTTT G TGATGTACA
 2701 GAGTGTATT ATTGACACGC CGGGCGACG GATGGTGATC CCCCTGGCCA GTGCACGTCT
 2761 GCTGTCAAGAT AAAGTCTCCC GTGAACTTTA CCCGGTGGTG CATATCGGGG ATGAAAGCTG
 2821 GCGCATGATG ACCACCGATA TGGCCAGTGT GCCGGTCTCC GTTATCGGGG AAGAAGTGGC
 2881 TGATCTCAGC CACCGCGAAA ATGACATCAA AAACCGCATT AACCTGATGT TCTGGGAAT
 2941 ATAAAATGTCA GGCTCCCTTA TACACAGCCA GTCTGCAGGT CGACCATAGT GACTGGATAT
 3001 GTTGTGTTT ACAGTATTAT GTAGTCTGTT TTTTATGAA AATCTAATT AATATATTGA
 3061 TATTTATATC ATTTTACGTT TCTCGTTCAG CTTCTTGTA CAAAGTGGTT TGATGGCCGC
 3121 TAAAGTAAAGTA AGACGTCGAG CTCTAAGTAA GTAACGGCCG CCACCGCGGT GGAGCTTGG
 3181 ACTTCTTCGC CAGAGGTTTG GTCAAGTCTC CAATCAAGGT TGTCGGCTTG TCTACCTTGC
 3241 CAGAAATTAA CGAAAAGATG GAAAAGGGTC AAATCGTTG TAGATACGTT GTTGACACTT
 3301 CTAAATAAGC GAATTCTTA TGATTTATGA TTTTATTAA TAAATAAGTT ATAAAAAAA
 3361 TAAGTGTATA CAAATTAAAGT AGTACTCTT AGGTTTTAAA ACGAAAATTTC TTATTCTGA
 3421 GTAATCTTT CCTGTAGGTC AGGTTGCTTT CTCAGGTATA GCATGAGGTC GCTCTTATTG
 3481 ACCACACCTC TACCGGCATG CCGAGCAAAT GCCTGCAAAT CGCTCCCCAT TTCACCCAAT
 3541 TGTAGATATG CTAACCTCCAG CAATGAGTTG ATGAATCTCG GTGTGTATT TATGTCTCA
 3601 GAGGACAATA CCTGTTGTA TCGTTCTCC ACACGGATCC CAATTCGCC TATAGTGAGT
 3661 CGTATTACAA TTCACTGGCC GTCGTTTAC AACGTCGTGA CTGGGAAAAC CCTGGCGTTA
 3721 CCCAACTTAA TCGCCCTGCA GCACATCCCC CTTTCGCCAG CTGGCGTAAT AGCGAAGAGG
 3781 CCCGCACCGA TCGCCCTTCC CAACAGTTGC GCAGCCTGAA TGGCGAATGG ACGGCCCTG
 3841 TAGCGCGCA TTAAGCGCGG CGGGTGTGGT GGTTACGCGC AGCGTGACCG CTACACTTGC
 3901 CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT CTTCCCTTCC TTTCTGCCA CGTTGCCGG
 3961 CTTTCCCCGT CAAGCTCTAA ATCGGGGGCT CCCTTTAGGG TTCCGATTTA GTGCTTACG
 4021 GCACCTCGAC CCCAAAAAAC TTGATTAGGG TGATGGTTCA CGTAGTGGGC CATGCCCTG
 4081 ATAGACGGTT TTTCGCCCTT TGACGTTGGA GTCCACGTT TTTAATAGTG GACTCTGTT
 4141 CCAAACCTGGA ACAACACTCA ACCCTATCTC GGTCTATTCT TTGATTTAT AAGGGATTTT
 4201 GCGGATTTG GCTTATTGGT TAAAAAAATGA GCTGATTTAA CAAAAAATTAA ACGCGAATTT
 4261 TAACAAAATA TTAACGTTA CAATTCCTG ATGCGGTATT TTCTCCCTAC GCATCTGTGC
 4321 GGTATTCAC ACCGCAGGCA AGTGCACAAA CAATACTTAA ATAAATACTA CTCAGTAATA
 4381 ACCTATTTCT TAGCATTTTT GACGAAATTT GCTATTTGT TAGAGTCTTT TACACCATT
 4441 GTCTCACAC CTCCGCTTAC ATCAACACCA ATAACGCCAT TTAATCTAAG CGCATCACCA
 4501 ACATTTCTG GCGTCAGTCC ACCAGCTAAC ATAAAATGTA AGCTTTCGGG GCTCTTGC
 4561 CTTCCAACCC AGTCAGAAAT CGAGTTCCAA TCCAAAAGTT CACCTGCTCC ACCTGCTTCT
 4621 GAATCAAACA AGGGAATAAA CGAATGAGGT TTCTGTGAAG CTGCACTGAG TAGTATGTTG
 4681 CAGTCTTTG GAAATACGAG TCTTTAATA ACTGGCAAAC CGAGGAATC TTGGTATTCT
 4741 TGCCACGACT CATCTCCATG CAGTTGGACG ATATCAATGC GTAACTATT GACCAGAGCC
 4801 AAAACATCCT CCTTAGGTTG ATTACGAAAC ACGCCAACCA AGTATTCGG AGTGCCTGAA
 4861 CTATTTTAT ATGCTTTTAC AAGACTTGA ATTTCCTTG CATAACGG GTCAATTGTT
 4921 CTCTTCTAT TGGGCACACA TATAATACCC AGCAAGTCAG CATCGGAATC TAGAGCACAT
 4981 TCTGGGGCCT CTGTGCTCTG CAAGCCGAA ACTTCACCA ATGGACAGGA ACTACCTGTG
 5041 AAATTAATAA CAGACATACT CCAAGCTGCC TTTGTGTGCT TAATCAGTA TACTCACGTG
 5101 CTCAATAGTC ACCAATGCC TCCCTCTGG CCCTCTCCCT TTCTTTTTC GACCGAATTA
 5161 ATTCTTAATC GGCAAAAAAA GAAAAGCTCC GGATCAAGAT TGTACGTAAG GTGACAAGCT
 5221 ATTTTCAAT AAAGAATATC TTCCACTACT GCCATCTGGC GTCATAACTG CAAAGTACAC
 5281 ATATATTACG ATGCTGTCTA TTAAATGCTT CCTATATTAT ATATATAGTA ATGCTTAA
 5341 TGGTGCACCTC TCAGTACAAT CTGCTCTGAT GCCGCATAGT TAAGCCAGCC CCGACACCCG
 5401 CCAACACCCG CTGACCGGCC CTGACGGGCT TGTCTGCTCC CGGCATCCGC TTACAGACAA
 5461 GCTGTGACCG TCTCCGGGAG CTGCATGTGT CAGAGGTTTT CACCGTCATC ACCGAAACGC
 5521 GCGAGACGAA AGGGCCTCGT GATAACGCTA TTTTATAGG TTAATGTCT GATAATAATG
 5581 GTTCTTAGG ACGGATCGCT TGCCGTAAAC TTACACGCC CTCGTATCTT TTAATGATGG
 5641 AATAATTTGG GAATTACTC TGTGTTTATT TATTTTATG TTTTGTATT GGATTTAGA
 5701 AAGTAAATAA AGAAGGTAGA AGAGTTACGG AATGAAGAAA AAAAAATAAA CAAAGGTTA
 5761 AAAAAATTCA ACACAAAGCG TACTTTACAT ATATATTTAT TAGACAAGAA AAGCAGATTA
 5821 AATAGATATA CATTGATTA ACGATAAGTA AAATGAAAAA TCACAGGATT TTGCTGTG
 5881 GTCTTCTACA CAGACAAGAT GAAACAATTG GGCATTAATA CCTGAGAGCA GGAAGAGCAA
 5941 GATAAAAGGT AGTATTTGTT GGCGATCCCC CTAGAGTCTT TTACATCTC GGAAAACAAA
 6001 AACTATTTT TCTTTAATTT CTTCTTAC TTTCTATT TAAATTTAT ATTTATATTA-

FIGURE 42c

119/240

6061 AAAAATTTAA ATTATAATT TTTTATAGC ACGTGATGAA AAGGACCCAG GTGGCACTTT
 6121 TCGGGGAAAT GTGCGCGAA CCCCTATTG TTTATTTTC TAAATACATT CAAATATGTA
 6181 TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA GGAAGAGTAT
 6241 GAGTATTCAA CATTCCGTG TGCGCCTTAT TCCCTTTTT GCGGCATTG GCCTTCCTGT
 6301 TTTTGCTCAC CCAGAAACGC TTGTTAAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG
 6361 AGTGGGTTAC ATCGAACTGG ATCTCAACAG CGGTAAGATC TTGAGAGTT TCGCCCCGA
 6421 AGAACGTTT CCAATGATGA GCACTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG
 6481 TATTGACGCC GGGCAAGAGC AACTCGGTG CGCATAACAC TATTCTCAGA ATGACTTGGT
 6541 TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG
 6601 CAGTGTGCC ATAACCATGA GTGATAACAC TGCGGCCAC TTACTTCTGA CAACGATCGG
 6661 AGGACCGAAG GAGCTAACCG CTTTTTTCA CAACATGGGG GATCATGTAA CTCGCCTTGA
 6721 TCGTTGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC
 6781 TGTAGCAATG GCAACAAACGT TGCAGCAACT ATTAACTGGC GAACTACTTA CTCTAGCTTC
 6841 CCGGCAACAA TTAATGACT GGATGGAGGC GGATAAAAGTT GCAGGACCAC TTCTGCGCTC
 6901 GGCCTTCCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCGGGTGAGC GTGGGTCTCG
 6961 CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC
 7021 GACGGGCACTG CAGGCAACTA TGAGATGAAAG AAATAGACAG ATCGCTGAGA TAGGTGCCTC
 7081 ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT AGATTGATT
 7141 AAAACTCAT TTTTAATTAA AAAGGATCTA GGTGAAGATC CTTTTGATA ATCTCATGAC
 7201 CAAAATCCCT TAACGTGAGT TTTCTGTTCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA
 7261 AGGATCTCT TGAGATCCTT TTTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC
 7321 ACCGCTACCA GCGGTGGTTT GTTTGCGGA TCAAGAGCTA CCAACTCTT TTCCGAAGGT
 7381 AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTCCTT CTAGTGTAGC CGTAGTTAGG
 7441 CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCCTGTTACC
 7501 AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAAGTT
 7561 ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGG
 7621 GCGAACGACC TACACCGAAC TGAGATACT ACAGCGTGAG CATTGAGAAA GCGCCACGCT
 7681 TCCCAGGGG AGAAAGCGG ACAGGTATCC GGTAAAGCGGC AGGGTCCGAA CAGGAGAGCG
 7741 CACGAGGGAG CTTCAGGGG GGAACGCTG GTATCTTAT AGTCTGTGCG GTTTTGC
 7801 CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGAGCC TATGGAAAAA
 7861 CGCCAGCAAC GCGGCCTTT TACGGTTCTT GGCCTTTGC TGGCTTTTG CTCACATGTT
 7921 CTTTCTGCG TTATCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA
 7981 TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTC GTGAGCGAGG AAGCGGAAGA
 8041 GCGCCAATA CGCAAACCGC CTCTCCCCGCG GCGTTGGCCG ATTCAATTAAAT GCAGCTGGCA
 8101 CGACAGGTTT CCGCACTGGA AAGCGGGCAG TGAGCGCAAC GCAATTAAAT TGAGTTACCT
 8161 CACTCATTAG GCACCCCGAG CTTTACACTT TATGCTTCCG GCTCTATGT TGTGTGGAAT
 8221 TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC CATGATTACG CCAAGCTCGG
 8281 AATTAACCCCT CACTAAAGGG AACAAAGCT GGGTACCGGG CCCCCCTCG AGATCCGGGA
 8341 TCGAAGAAAT GATGGTAAAT GAAATAGGAA ATCAAGGAGC ATGAAGGCAA AAGACAAATA
 8401 TAAGGGTCGA ACGAAAATA AAGTAAAAG TGTTGATATG ATGTATTG TGTTGGCGCG
 8461 CCGAAAAAAC GAGTTTACCG AATTGCAACAA TCATGCTGAC TCTGTGGCGG ACCCGCGCTC
 8521 TTGCCGGCCC GGCGATAACG CTGGGCGTGA GGCTGTGCC GGCGGAGTT TTGCGCCTG
 8581 CATTITCCAA GTTTTACCT GCGCTAAGGG GCGAGATTGG AGAAGCAATA AGAATGCC
 8641 TTGGGGTTGC GATGATGAGC ACCACGACAA CTGGTGTAT TATTTAAGTT GCCGAAAGAA
 8701 CCTGAGTGCA TTGCAACAT GAGTACTA GAAGAATGAG CCAAGACTTG CGAGACGCGA
 8761 GTTTGCCGGT GGTGCGAACATA GAGACGAC CATGACCTTG AAGGTGAGAC GCGCATAACC
 8821 GCTAGAGTAC TTTGAAGAGG AAACAGCAAT AGGGTTGCTA CCAGTATAAA TAGACAGGTA
 8881 CATACAACAC TGGAAATGGT TGTCTGTTG AGTACGCTTT CAA

FIGURE 420

120/240

pDEST23

His6 carboxy-fusion vector, T7 promoter

T7 Promoter → mRNA
 205 atc ccg cga aat taa tac gac tca cta tag ggc gat cac aac ggt ttc cct
 tag ggc gct tta att atg ctg agt gat atc cgt ctg gtg ttg cca aag gga
 256 cta gat cac aag ttt gta caa aaa aac tga acg aca gta aaa tga tat //
 gat cta ggg ttc aaa cat gtt ttt tcc act tcc tat ttg cat ttt act ata //

// — CmR — ccdB — //

1888 ttt tta tgc aaa ato taa ttt aat ata ttg ata ttt ata tca ttt tac gtt
 aaa aat acg ttt tag att aaa tta tat sac tat aaa tat agt aaa atg caa
 attR2 A F L Y K V Y I M S Y Y H H
 1939 tct cgt tca gct ttd ttg tac aaa gtg gtg att atg tcc tac tac cat cac
 aga gca agt cga aag aac atg ttt oac cac taa tac aac atg atg gta gta //
 His6 //
 1990 cat cat cat cac ctc gat gag caa taa cta gca taa ccc ctt ggg goc tot
 gta gta gta gta gat ctc gtt att gat cgt att ggg gaa ccc cgg aga

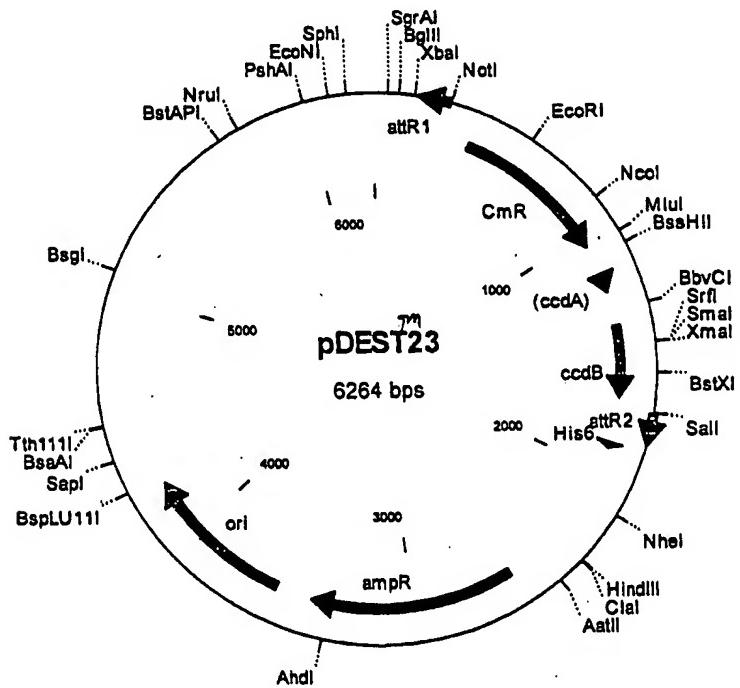


FIGURE 43A

121/240

pDEST23 6264 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
285..161	attR1
394..1053	CmR
1173..1257	inactivated ccdA
1395..1700	ccdB
1741..1865	attR2
1883..1911	his6
2574..3434	ampR
3583..4222	ori

1 TCTTCCCCAT CGGTGATGTC GGCGATATA GCGCCAGCAA CCGCACCTGT GGCGCCGGTG
 61 ATGCCGCCA CGATGCGTCC GGCCTAGAGG ATCGAGATCT CGATCCCGCG AAATTAATAC
 121 GACTCACTAT AGGGAGACCA CAACGGTTTC CCTCTAGATC ACAAGTTTGT ACAAAAAAAGC
 181 TGAACGAGAA ACGTAAAATG ATATAAATAT CAATATATTA AATTAGATTT TGCATAAAAA
 241 ACAGACTACA TAATACTGTA AAACACAACA TATCCAGTCA CTATGGCGGC CGCATTAGGC
 301 ACCCCAGGCT TTACACTTTA TGCTTCCGGC TCGTATAATG TGTGGATTTT GAGTTAGGAT
 361 CCGGCGAGAT TTTCAGGAGC TAAGGAAGCT AAAATGGAGA AAAAAATCAC TGGATATACC
 421 ACCGTTGATA TATCCCATTG GCATCGTAA GAACATTTG AGGCATTTC GTCAGTTGCT
 481 CAATGTACCT ATAACCAGAC CGTCAGCTG GATATTACGG CCTTTTTAAA GACCGTAAAG
 541 AAAAATAAGC ACAAGTTTA TCCGGCTTT ATTACACATTC TTGCCCCCCT GATGAATGCT
 601 CATCCGAAAT TCCGTATGGC AATGAAAGAC GGTGAGCTGG TGATATGGGA TAGTGTTCAC
 661 CCTTGTTCACA CCGTTTTCCA TGAGCAAAC GAAACATTTT CATCGCTCTG GAGTGAATAC
 721 CACGACGATT TCCGGCAGTT TCTACACATA TATTGCAAG ATGTGGCGTG TTACGGTGAA
 781 AACCTGCCCT ATTCCCTAA AGGGTTTATT GAGAATATGT TTTTGTCTC AGCCAATCCC
 841 TGGGTGAGTT TCACCAGTT TGATTTAAC GTGGCCAATA TGACAAACTT CTTCCGCCCC
 901 GTTTTCACCA TGGCAAATA TTATACGCAA GGCGACAAGG TGCTGATGCC GCTGGCGATT
 961 CAGGTTCATC ATGCCGTCTG TGATGGCTTC CATGTCGGCA GAATGCTTAA TGAATTACAA
 1021 CAGTACTGCG ATGAGTGGCA GGGCGGGGCG TAAACGGCTG GATCCGGCTT ACTAAAAGCC
 1081 AGATAACAGT ATGCGTATTG CGCGCTGAT TTTTGGGTA TAAGAATATA TACTGATATG
 1141 TATACCCGAA GTATGTCAA AAGAGGTGTG CTATGAAGCA GCGTATTACA GTGACAGTTG
 1201 ACAGCGACAG CTATCAGTTG CTCAAGGCAT ATATGATGTC AATATCTCCG GTCTGGTAAG
 1261 CACAACCATG CAGAATGAAG CCCGTCGTCT GCGTGGCGAA CGCTGGAAAG CGGAAATCA
 1321 GGAAGGGATG GCTGAGGTGCG CCCGGTTTAT TGAAATGAAAC GGCTTTTG CTGACGGAGAA
 1381 CAGGGACTGG TGAAATGCG TTTAAGGTTT ACACCTATAA AAGAGAGAGC CGTTATCGTC
 1441 TGTGGTGGAA TGTACAGAGT GATATTATTG ACACGCCGG GCGACGGATG GTGATCCCC
 1501 TGGCCAGTGC ACGTCTGCTG TCAGATAAAG TCTCCGTGA ACTTTACCCG GTGGTGCATA
 1561 TCGGGGATGA AAGCTGGCGC ATGATGACCA CCGATATGGC CAGTGTGGCG GTCTCCGTTA
 1621 TCGGGGAAAGA AGTGGCTGAT CTCAGCCACC GCGAAATGA CATCAAAAC GCCATTAACC
 1681 TGATGTTCTG GGGAAATATAA ATGTCAGGCT CCCTTATACA CAGCCAGTCT GCAGGTCGAC
 1741 CATAGTGACT GGATATGTTG TGTGTTACAG TATTATGTAG TCTGTTTTT ATGCAAATC
 1801 TAATTTATAA TATTGATATT TATATCATTT TACGTTTCTC GTTCAGCTTT CTIGTACAAA
 1861 GTGGTGATTA TGTCGACTA CCATCACCAT CACCATCACC TCGATGAGCA ATAACTAGCA
 1921 TAACCCCTTG GGGCCTCTAA ACGGGTCTG AGGGGTTTT TGCTGAAAGG AGGAACATATA
 1981 TCCGGATATC CACAGGACGG GTGGTGTGCG CATGATGCGC TAGTCGATAG TGGCTCCAAG
 2041 TAGCGAAGCG AGCAGGACTG GGGGGGGGAA AGCGGTGCG GACAGTGTCTC CGAGAACGGG
 2101 TGCGCATAGA AATTGCACTA ACGCATATAG CGCTAGCAGC ACGCCATAGT GACTGGCGAT
 2161 GCTGTCGAA TGGACGATAT CCCGCAAGAG GCCCAGCAGT ACCGGCATAA CCAAGCCTAT
 2221 GCCTACAGCA TCCAGGGTGA CGGTGCCGAG GATGACGATG AGCGCATTGT TAGATTCAT
 2281 ACACGGTGCC TGACTGCGTT AGCAATTAA CTGTGATAAA CTACCGCATT AAAGCTTATC
 2341 GATGATAAGC TGTCAAACAT GAGAATTCTT GAAGAAGAAA GGGCCTCGTG ATACGCTAT
 2401 TTTTATAGGT TAATGTCATG ATAATAATGG TTTCTTAGAC GTCAGGTGGC ACTTTTCGGG
 2461 GAAATGTGCG CGGAACCCCT ATTTGTTTAT TTTTCTAAAT ACATTCAAAT ATGTATCCGC
 2521 TCATGAGACA ATAACCCCTGA TAAATGCTTC AATAATATTG AAAAGGAAG AGTATGAGTA
 2581 TTCAACATTT CCGTGTGCGC CTTATTCCCT TTTTGGCGC ATTTTGCCTT CCTGTTTTG
 2641 CTCACCCAGA AACGCTGGTG AAAGTAAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG-

FIGURE 43B

122/240

2701 GTTACATCGA ACTGGATCTC AACAGCGGT AAGTCCTGA GAGTTTCGC CCCGAAGAAC
 2761 GTTTCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG CGCGGTATTA TCCCCTGTTG
 2821 ACGCCGGCA AGAGCAACTC GGTCGCCGA TACACTATT TCAGAATGAC TTGGTTGAGT
 2881 ACTCACCACT CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGCAGTG
 2941 CTGCCATAAC CATGAGTGT AACACTGCGG CCAACTTACT TCTGACAACG ATCGGAGGAC
 3001 CGAAGGAGCT AACCGTTTT TTGACAACA TGGGGGATCA TGAACTCGC CTTGATCGTT
 3061 GGGAACCGGA GCTGAATGAA GCATACCAA ACGACGAGCG TGACACCACG ATGCCTGCAG
 3121 CAATGGCAAC AACGTTGCGC AAACATTTAA CTGGCGAAT ACCTACTCTA GCTTCCCCGC
 3181 AACAAATAAT AGACTGGATG GAGGCGGATA AAGTTGCAAGG ACCACCTCTG CGCTCGGGCC
 3241 TTCCGGCTGG CTGGTTTATT GCTGATAAT CTGGAGCCGG TGAGCGTGGG TCTCGCGTA
 3301 TCATTGCGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG
 3361 GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA
 3421 TTAAGCATTG GTAATGTC GACCAAGTTT ACTCATATAT ACCTTAGATT GATTAAAAC
 3481 TTCATTTTA ATTAAAAGG ATCTAGGTGA AGATCCTTT TGATAATCTC ATGACCAAA
 3541 TCCCTTAACG TGAGTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT
 3601 CTTCTTGAGA TCCTTTTTT CTGCGCGTAA TCTGCTGCTT GCAAACAAAA AAACCACCGC
 3661 TACCAGCGGT GGTTTGTGG CGGATCAAG AGCTACCAAC TCTTTTCCG AAGGTAACGT
 3721 GCTTCAGCAG AGGCAGATA CCAAATACTG TCCCTCTAGT GTAGCGTAG TTAGGCCACC
 3781 ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAATCCTG TTACCAAGTGG
 3841 CTGCTGCCAG TGGCGATAAG TCGTGTCTT CCGGGTTGGA CTCAAGACGA TAGTTACCGG
 3901 ATAAGGCAGCA GCGGTGGGC TGAAACGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA
 3961 CGACCTACAC CGAATGAGA TACCTACAGC GTGAGCTATG AGAAAGCGCC ACGCTTCCCG
 4021 AAGGGAGAAA GGCGGACAGG TATCCGGTAA CGGGCAGGGT CGGAACAGGA GAGCGCACGA
 4081 GGGAGCTTCC AGGGGGAAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT CGCCACCTCT
 4141 GACTTGAGCG TCGATTTTG TGATGCTCGT CAGGGGGGG GAGCCTATGG AAAAACGCCA
 4201 GCAACGCGGC CTTTTTACGG TTCCCTGGCTT TTTGCTGGCC TTTTGTCTAC ATGTTCTTC
 4261 CTGCGTTATC CCGTGATTCT GTGGATAACC GTATTACCGC TTGAGTGA GCTGATAACCG
 4321 CTCGCCGCAG CGGAACGACC GAGCGCAGCG AGTCAGTGG CGAGGAAGCG GAAAGAGCGC
 4381 TGATGCGGT TTTCTCCTT ACGCATCTGT CGGGTATTT ACACCGCATA TATGGTGCAC
 4441 TCTCAGTACA ATCTGCTCTG ATGCCGATA GTTAAGCCAG TATACACTCC GCTATCGCTA
 4501 CGTGACTGGG TCATGGCTGC GCCCCGACAC CGGCCAACAC CGCTGACGC GCCCTGACGG
 4561 GCTTGCTGC TCCCGGCATC CGCTTACAGA CAAGCTGTGA CGCTCTCCGG GAGCTGCATG
 4621 TGTCAGGAGT TTTCACCGTC ATCACCGAAA CGCGCAGGGC AGCTCGGGTA AAGCTCATCA
 4681 GCGTGGTCGT GAAGCGATTG ACAGATGTC GCCTGTTCAT CGCGTCCAG CTCGTTGAGT
 4741 TTCTCCAGAA GCGTTAATGT CTGGCTTCTG ATAAAGCGGG CCATGTTAAG GGCCTTTT
 4801 TCCCTGTTGG TCACTGATGC CTCCGTGAA GGGGGATTT TGTTCATGGG GGTAATGATA
 4861 CCGATGAAAC GAGAGAGGAT GCTCACGATA CGGGTTACTG ATGATGAACA TGCCCCGTTA
 4921 CTGGAACGTT GTGAGGGTAA ACAACTGGCG GTATGGATGC GGCGGGACCA GAGAAAATC
 4981 ACTCAGGGTC AATGCCAGCG CTGGCTTAAAT ACAGATGTCG GTGTTCCACA GGGTAGCCAG
 5041 CAGCATCTG CGATGCGAGT CGGGAAACAT ATGGTGCAGG GCGCTGACTT CCGCGTTTCC
 5101 AGACTTTACG AAACACGGAA ACCGAAGACC ATTCACTGTT TGCTCAGGT CGCAGACGTT
 5161 TTGCAGCAGC AGTCGCTTC CGTTCGCTCG CGTATCGGTG ATTCAATTCTG CTAACCGTA
 5221 AGGCACCCCCC GCCACCTCG CCGGGTCTTC AACGACAGGA CGACGATCAT GCGCACCCGT
 5281 GGCCAGGACC CAACGCTGCC CGAGATGCC CGCGTGCAGG TGCTGGAGAT GCGGGACGCG
 5341 ATGGATATGT TCTGCCAAGG GTGGTTTGC GCATTACAG TTCTCCGAA GAATTGATTG
 5401 GCTCCAATTC TTGGAGTGGT GATCCTGTA CGCAGGTGCC GCGGGCTTC ATTCAAGGTCG
 5461 AGGTGGCCCG GCTCCATGCA CGCGACGCCA CGCGGGGAG CGAGACAAGG TATAGGGCG
 5521 CGCCTACAAT CCATGCCAAC CGCTTCCATG TGCTCGCGA GCGGGCATAA ATCGCCGTGA
 5581 CGATCAGCGG TCCAGTGATC GAAGTTAGGC TGGTAAGAGC CGCGAGCGAT CCTTGAAGCT
 5641 GTCCCTGATG GTCGTCATCT ACCTGCCTGG ACAGCATGGC CTGCAACGCG GGCATCCGA
 5701 TGCCGCCGA AGCGAGAAGA ATCATTAATGG GGAAGGCCAT CCAGCCTCGC GTCGCGAACG
 5761 CCAGCAAGAC GTAGCCCAGC GCGTCGGCCG CCATGCCGGC GATAATGGCC TGCTTCTCGC
 5821 CGAAACGTTT GTGGCGGGG CGAGTGCAGA AGGCTTGAGC GAGGGCGTGC AAGATTCCGA
 5881 ATACCGCAAG CGACAGGCCG ATCATCGTC CGCTCCAGCG AAAGCGGTCC TCGCCGAAAA
 5941 TGACCCAGAG CGCTGCCGGC ACCTGTCCTA CGAGTTGCAT GATAAGAAG ACAGTCATAA
 6001 GTGCGGCAC GATAGTCATG CCCCGCAGCC ACCGGAAGGA GCTGACTGGG TTGAAGGCTC
 6061 TCAAGGGCAT CGGTGCGATCG ACGCTCTCCC TTATGCGACT CCTGCAATTAG GAAGCAGCCC
 6121 AGTAGTAGGT TGAGGCCGTT GAGCACCAGG GCCGCAAGGA ATGGTGCATG CAAGGAGATG-

FIGURE 43C

123/240

6181 GCGCCCAACA GTCCCCGGC CACGGGCCT GCCACCATA CCACGCCGAA ACAAGCGCTC
6241 ATGAGCCCGA AGTGGCGAGC CCGA

FIGURE 43D

174/260

pDEST24
GST carboxy-fusion vector, T7 promoter

T7 Promoter → mRNA

```

1  atc gag atc tcc atc ccc ega aat taa tac gac tca cta tag gaa gac ccc
tag ctc tag agc tag ggc get tta att atg ctg agt gat atc ccc ctg gtg
      att R1
      att R2
52  aac ggt ttc oct cta gat cac aag ttt gta cca aaa agc tga acg aga aac
ttg cca aag gga gat cta ggg ttc aaa cat gtt ttg tcg act tgc tct ttg

```

// — CmR — ccdB — //

att R2 A F L Y K V V I M S

```

1725 // tca tcc tcc gtt tct egt tca gct tcc ttg tac aaa gtc gtb att atg tcc
agt aaa atg cca aca qca agt cga aag abc atq ttc ccc ccc taa tac agg
      P I L      GST Protein → (~ 225 kDa)
1786 // oct ata cta ggt tat tgg aaa att aag ggc ttt gtg ccc act cga ctt
gga tat gat cca ata acc ttt taa ttc ccc gaa cac gtt ggg tga gct gaa

```

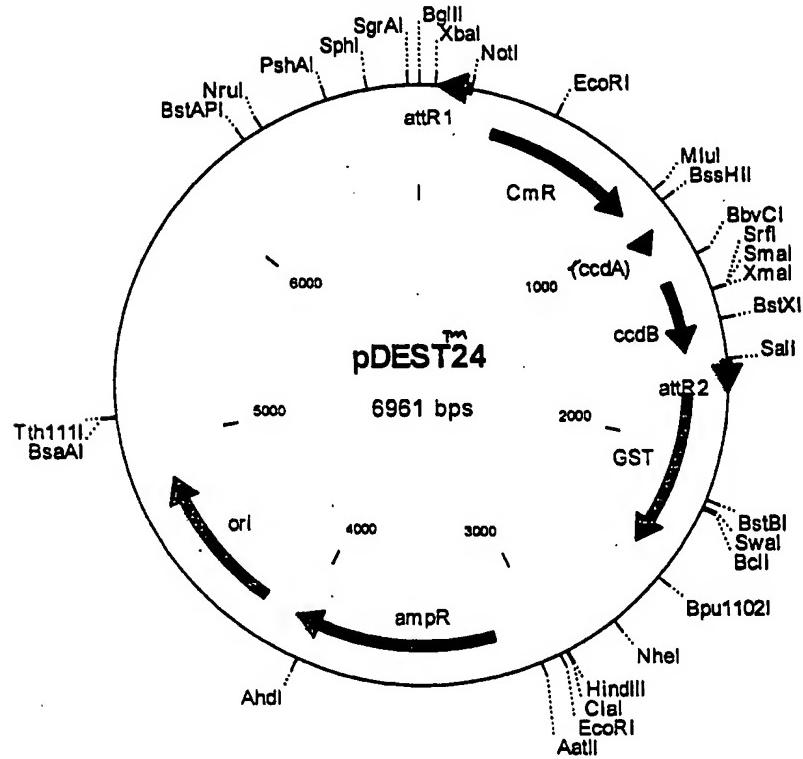


FIGURE 44A

125/240

pDEST24 6961 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
195..71	attR1
304..963	CmR
1083..1167	inactivated ccda
1305..1510	ccdB
1651..1775	attR2
1783..2451	GST
3181..4041	ampR
4190..4829	ori

1 ATCGAGATCT CGATCCCGCG AAATTAATAC GACTCACTAT AGGGAGACCA CAACGGTTTC
 61 CCTCTAGATC ACAAGTTTGT ACAAAAAAGC TGAAACGAGAA ACGTAAAATG ATATAAATAT
 121 CAATATATTA AATTAGATTT TGCATAAAAA ACAGACTACA TAATACTGTA AAACACAACA
 181 TATCCAGTCA CTATGGCGGC CGCATTAGGC ACCCAGGCT TTACACTTTA TGCTTCGGC
 241 TCGTATAATG TGTGGATTTT GAGTTAGGAT CCGGCGAGAT TTTCAGGAGC TAAGGAAGCT
 301 AAAATGGAGA AAAAATCAC TGGATATACC ACCGTTGATA TATCCCAATG GCATCGAAA
 361 GAACATTTG AGGCATTCA GTCAGTTGCT CAATGTACCT ATAACCAGAC CGTTCAAGCTG
 421 GATATTACGG CCTTTTTAAA GACCGTAAAG AAAAATAAGC ACAAGTTTA TCCGGCCTTT
 481 ATTACACATTC TTGCCCCGCT GATGAATGCT CATCCGGAAT TCCGTATGGC AATGAAAGAC
 541 GGTGAGCTGG TGATATGGGA TAGTGTTCAC CCTTGTACCA CGTTTCCCA TGAGCAAAC
 601 GAAACTTT CATCGCTCTG GAGTGAATAC CACGACGATT TCCGGCAGTT TCTACACATA
 661 TATTGCAAG ATGTGGCGTG TTACGGTGA AACCTGGCCT ATTTCCCTAA AGGGTTTATT
 721 GAGAATATGT TTTTCGTCTC AGCCAATCCC TGGGTGAGTT TCACCAAGTT TGATTTAAC
 781 GTGGCCAATA TGGACAACCTT CTTCGCCCTT GTTTCACCA TGGGCAAATA TTATACGCAA
 841 GGCGACAAGG TGTGTATGCC GCTGGCGATT CAGGTTCATC ATGCCGTCTG TGATGGCTTC
 901 CATGTCGGCA GAATGCTTAA TGAATTACAA CAGTACTGCG ATGAGTGGCA GGGCGGGGCG
 961 TAAACGCGTG GATCCGGCTT ACTAAAAGCC AGATAACAGT ATGCGTATTG GCGCGCTGAT
 1021 TTTTGGGTA TAAGAATATA TACTGATATG TATACCCGAA GTATGTCAAA AAGAGGTGTG
 1081 CTATGAAAGCA GCGTATTACA GTGACAGTTG ACAGCGACAG CTATCAGTTG CTCAAGGCAT
 1141 ATATGATGTC AATATCTCCG GTCTGGTAAAG CACRAACCATG CAGAATGAAG CCCGTCGTCT
 1201 GCGTGGCGAA CGCTGGAAAG CGGAAAATCA GGAAGGGATG GCTGAGGTCG CCCGGTTTAT
 1261 TGAAATGAAC GGCTCTTTG CTGACGAGAA CAGGGACTGG TGAAATGCAAG TTTAAGGTTT
 1321 ACACCTATAA AAGAGAGAGC CGTTATGTC TGTTTGTGGA TGTACAGAGT GATATTATTG
 1381 ACACGCCCGG GCGACGGATG GTGATCCCCC TGGCAGTGC ACGCTGCTG TCAGATAAAG
 1441 TCTCCCGTGA ACTTTACCCG GTGGTGCATA TCGGGGATG AAGCTGGCGC ATGATGACCA
 1501 CCGATATGGC CAGTGTGCCG GTCTCCGTTA TCGGGGAAAGA AGTGGCTGAT CTCAGCCACC
 1561 GCGAAAATGA CATCAAAACGCC ATGATTTCTG GGGAAATATAA ATGTCAGGCT
 1621 CCCTTATACA CAGCCAGTCT GCAGGTCGAC CATAGTGAAT GGATATGTTG TGTTTACAG
 1681 TATTATGTAG TCTGTTTTT ATGCAAAATC TAATTAAATA TATTGATATT TATATCATTT
 1741 TACGTTCTC GTTCACTTTT CTTGTACAA GTGGTGATTA TGTCCTCTAT ACTAGGTTAT
 1801 TGGAAAATTA AGGGCTTGTG GCAACCCACT CGACTCTTT TGGAATATCT TGAAAGAAAA
 1861 TATGAAGAGC ATTTGTATGA GCGCGATGAA GGTGATAAT GGCAGAAACAA AAAGTTGAA
 1921 TTGGGTTTGG AGTTTCCCA TCTTCCTTAT TATATTGATG GTGATGTTAA ATTAACACAG
 1981 TCTATGGCCA TCATACGTTA TATAGCTGAC AAGCACAAACA TGTTGGGTTG TTGTCCAAA
 2041 GAGCGTGCAG AGATTTCAAT GCTTGAAGGA CGGGTTTGG ATATTAGATA CGGTGTTTCG
 2101 AGAATTGCAT ATAGTAAAGA CTTTGAACAT CTCAAAGTTG ATTTTCCTAG CAAGCTACCT
 2161 GAAATGCTGA AAATGTTGCA AGATCGTTA TGTCATAAAA CATATTTAAA TGGTGATCAT
 2221 GTAAACCCATC CTGACTTCAT GTGGTATGAC GCTCTTGATG TTGTTTTATA CATGGACCCA
 2281 ATGTCCTGG ATGCGTTCCC AAAATTAGTT TGTTTTAAA AACGTATTGA AGCTATCCCA
 2341 CAAATTGATA AGTACTTGAATCCAGCAAG TATATAGCAT GGCCTTTGCA GGGCTGGCAA
 2401 GCCACGTTTG GTGGTGGCGA CCATCCTCCA AAATCGGATC TGTTCCGGC TCCATGGGGA
 2461 TCCGGCTGCT AACAAAGCCC GAAAGGAAGC TGAGTTGGCT GCTGCCACCG CTGAGCAATA
 2521 ACTAGCATAA CCCCTTGGGG CCTCTAAACG GGTTTGAGG GGTTTTITGC TGAAAGGAGG
 2581 AACTATATCC GGATATCCAC AGGACGGGTG TGGTCGCCAT GATCGCGTAG TCGATAGTGG
 2641 CTCCAAGTAG CGAAGCGAGC AGGACTGGC GGGGGCCAAA GCGGTGGAC AGTGCCTCGA-

FIGURE 44B

126/240

2701 GAACGGGTGC GCATAGAAAT TGCATCAACG CATATAGCGC TAGCAGCACG CCATAGTGAC
 2761 TGGCGATGCT GTCGGAATGG ACGATATCCC GCAAGAGGCC CGGCAGTACC GGCATAACCA
 2821 AGCCTATGCC TACAGCATCC AGGGTGACGG TGCCGAGGAT GACGATGAGC GCATTGTTAG
 2881 ATTTCATACA CGGTGCTGA CTGCGTTAGC AATTTAACTG TGATAAAACTA CCGCATTA
 2941 GCTTATCGAT GATAAGCTGT CAAACATGAG AATTCTTGAA GACGAAAGGG CCTCGTGATA
 3001 CGCCTATTT TATAGTTAA TGTATGATA ATAATGGTTT TTAGACGTC AGGTGGCACT
 3061 TTTCGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG
 3121 TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTCAAT AATATTGAAA AAGGAAGAGT
 3181 ATGAGTATTC AACATTTCCG TGCGCCCCC ATTCCCTTT TTGCGGCATT TTGCGCTTCC
 3241 GTTTTGCTC ACCCAGAAAC GCTGGTGAAGA GTAAAAGATG CTGAAGATCA GTTGGGTGCA
 3301 CGAGTGGGTT ACATCGAACT GGATCTAAC AGCGGTAAGA TCCTTGAGAG TTTTCGGCCCC
 3361 GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC
 3421 CGTGGTGCAG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG
 3481 GTTGAGTACT CACCAAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA
 3541 TGCACTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC
 3601 GGAGGACCGA AGGAGCTAAC CGCTTTTTG CACAACATGG GGGATCATGT AACTCGCC
 3661 GATCGTTGGG AACCCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG
 3721 CCTGCAGCAA TGGCAACAACT GTTGCAGCAA CTATTAACCTG GCGAAGTACT TACTCTAGCT
 3781 TCCCAGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAAG TTGCAAGGACC ACTTCTGC
 3841 TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAAATCTG GAGCCGGTGA GCGTGGGTCT
 3901 CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC
 3961 ACGACGGGGA GTCAAGCAAC TATGGATGAA CGAAATAGAC AGATCGTGA GATAGGTGCC
 4021 TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTACT CATATATACT TTAGATTGAT
 4081 TAAAAACTTC ATTTTTAATT TAAAAGGATC TAGGTGAAGA TCCTTTTGTA TAATCTCATG
 4141 ACCAAAATCC CTTAACGTGA GTTTTCGTT CACTGAGCGT CAGACCCCGT AGAAAAGATC
 4201 AAAGGATCTT CTTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA
 4261 CCACCGCTAC CAGCGGTGGT TTGTTGCGG GATCAAGAGC TACCAACTCT TTTCCGAAG
 4321 GTAACGGCT TCAGCAGAGC GCAGATAACCA AATACTGTCC TTCTAGTGT GCGTAGTTA
 4381 GGCCACCACT TCAAGAACTC TGTTAGCACCG CCTACATACC TCGCTCTGCT AATCCTGTTA
 4441 CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG GTTGGACTC AAGACGATAG
 4501 TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGTT CGTGCACACA GCCCAGCTTG
 4561 GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGT AGCTATGAGA AAGCGCCACG
 4621 CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CGGTTAAGCG GCAGGGTCGG AACAGGAGAG
 4681 CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT CGGGTTTC
 4741 CACCTCTGAC TTGAGCGTCG ATTTTTGTA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA
 4801 AACGCCAGCA ACGCGGCCCTT TTACGGTTC CTGGCCTTT GCTGGCCCTT TGCTCACATG
 4861 TTCTTCTCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCCTT TGAGTGAGCT
 4921 GATACCGCTC GCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA
 4981 GAGCGCTGA TGGGTATTT TCTCTTACG CATCTGTGCG GTATTTACA CCGCATATAT
 5041 GGTGCACCTCT CAGTACAATC TGCTCTGATG CCGCATAGTT AAGCCAGTAT AACTCCGCT
 5101 ATCGCTACGT GACTGGGTCA TGGCTGCGC CCGACACCCG CCAACACCCG CTGACGCC
 5161 CTGACGGGCT TGCTGCTCC CGGCATCCGC TTACAGACAA GCTGTGACCG TCTCCGGGAG
 5221 CTGCATGTG CAGAGTTT CACCGTCATC ACCGAAACGC GCGAGGGCAGC TGCAGTAAAG
 5281 CTCATCAGCG TGGCTGTGAA GCGATTACA GATGCTGCGC TTGTCATCCG CGTCCAGCTC
 5341 GTTGAGTTTC TCCAGAAGCG TTAATGTCTG GCTTCTGATA AAGCGGGCCA TGTTAAGGGC
 5401 GGTTTTTCC TTGTTGGTCA CTGATGCCCT CGTGTAAAGGG GGATTTCTGT TCATGGGGT
 5461 AATGATACCG ATGAAACGAG AGAGGATGCT CACGATACGG GTTACTGATG ATGAACATGC
 5521 CGGGTTACTG GAACGTTGTG AGGGTAAACA ACTGGCGGTG TGATGCGGGC GGGACCAAGAG
 5581 AAAAATCACT CAGGGTCAAT GCCAGCGCTT CGTTAATACA GATGTAGGTG TTCCACAGGG
 5641 TAGCCAGCAG CATCCTGCGA TGCAAGTCGG GAACATAATG GTGCAGGGCG CTGACTTCC
 5701 CGTTTCCAGA CTTTACGAAA CACGGAAACCGA GAAGACCGATT CATGTTGTTG CTCAGGTGCG
 5761 AGACGTTTTG CAGCAGCAGT CGCTTCACGT TCGCTCGCGT ATCGGTGATT CATTCTGCTA
 5821 ACCAGTAAGG CAACCCCGCC AGCCTAGCCG GGTCTCTAAC GACAGGAGCA CGATCATGCG
 5881 CACCCGTGGC CAGGACCCAA CGCTGCCGA GATGCGCCGC GTGCGGCTGC TGGAGATGGC
 5941 GGACGCGATG GATATGTTCT GCCAAGGGTT GGTTTGCAGCA TTACAGTTTC TCCGCAAGAA
 6001 TTGATTGGCT CCAATTCTG GAGTGGTGAA TCCGTTAGCG AGGTGCGGCC GGCTTCCATT
 6061 CAGGTGAGG TGGCCGGCT CCATGCACCG CGACGCAACG CGGGGAGGCA GACAAGGTAT
 6121 AGGGCGCGC CTACAATCCA TGCCAACCG TTCCATGTGC TCGCCGAGGC GGCATAAATC-

FIGURE 44C

127/240

6181 GCCGTGACGA TCAGCGGTCC AGTGATCGAA GTTAGGCTGG TAAGAGCCGC GAGCGATCCT
6241 TGAAGCTGTC CCTGATGGTC GTCATCTACC TGCTGGACA GCATGGCCTG CAACGCGGC
6301 ATCCCGATGC CGCCGGAAGC GAGAAGAAC ATAATGGGA AGGCCATCCA GCCTCGCGTC
6361 GCGAACGCCA GCAAGACGTA GCCCAGCGCG TCGGCCGCCA TGCCGGCGAT AATGGCCTGC
6421 TTCTCGCCGA AACGTTTGGT GGCGGGACCA GTGACGAAGG CTTGAGCGAG GGCAGTCAAG
6481 ATTCCGAATA CCGCAAGCGA CAGGCCGATC ATCGTCGCGC TCCAGCGAAA GCGGTCCCTCG
6541 CCGAAAATGA CCCAGAGCGC TGCCGGCACC TGTCCCTACGA GTGACATGAT AAAGAAGAC
6601 GTCATAAGTG CGCGACGAT AGTCATGCC CGCGCCCACC GGAAGGGAGCT GACTGGGTTG
6661 AAGGCTCTCA AGGGCATCGG TCGATCGACG CTCTCCCTTA TGCGACTCCT GCATTAGGAA
6721 GCAGCCCCAGT AGTAGGTTGA GGCGGTTGAG CACCGCCGCC GCAAGGAATG GTGACATGCAA
6781 GGAGATGGCG CCCAACAGTC CCCCAGGCCAC GGGGCTGCC ACCATACCCA CGCCGAAACA
6841 AGCGCTCATG AGCCCAGGT GGCGAGCCCG ATCTTCCCCA TCGGTGATGT CGGCGATATA
6901 GGCGCCAGCA ACCGCACCTG TGGCGCCGGT GATGCCGGCC ACGATGCGTC CGGCGTAGAG
6961 G

FIGURE 4D

128/240
FIGURE 45A

pDEST25
Thioredoxin carboxy-fusion vector, T7 promoter

T7 Promoter → mRNA

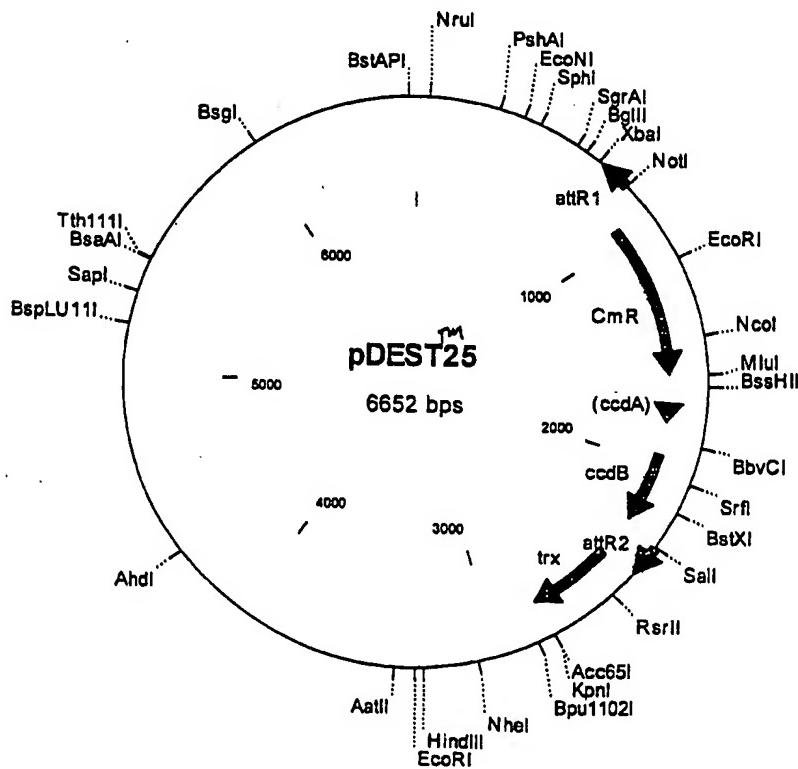
1 nag atc tcc atc ccc cga aat tta tac gao tca cta tay yaa yac cxc aac
 ntc tag agc tag ggc gct tta att atg ctg agt gat aac cat ctg gtg ttg

52 ggt ttc cct cta gat cac aag ttt gtt caa aaa agc tga acg aga aac gta
cca aag gga gat cta gca ttc aaa cat gtt ttg act tgg tct ttg cat

1 — CmR — ccdB — //

1735 // attR2 — A F₁ S_L Y K V V I M S D
 ttt tac gtt tct cgt tca get ttc ttg tac aaa gtc gtc att atg ago gat
 aaa atg caa aga gca agt cga aag aac atg ttt cac ccc taa tac tcc cta

1786 K I I — Trx Protein (~120 aa.) →
 aaa att att cac ctg act gac gac agt ttt gac atg gat gta ctc aaa gcg
 ttt taa taa gtc gac tga ctg ctg tca aaa ctg tgc cta cat gag ttt cgc



pDEST25 6652 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
844..720	attR1
953..1612	CmR
1732..1816	inactivated ccdA
1954..2259	ccdB
2300..2424	attR2
2432..2794	trx

1 CCGGAAGCGA GAAGAACAT AATGGGAAAG GCCATCCAGC CTCGCCTCGC GAACGCCAGC
 61 AAGACGTAGC CCAGCGCTC GGCGCCATG CGGGCGATAA TGGCCTGCTT CTCGCCAAA
 121 CGTTGGTGG CGGGACCAGT GACGAAGGCT TGAGCGAGGG CGTGCAAGAT TCCGAATACC
 181 GCAAGCGACA GGCGATCAT CGTCGCCTC CAGCGAAAGC GGTCCTCGCC GAAAATGACC
 241 CAGAGCGCTG CGGGCACCTG TCCTACGAGT TGCATGATAA AGAACAGACT CATAAGTGCG
 301 GCGACGATAG TCATGCCCG CGCCCACCGG AAGGGACTGA CTGGGTTGAA GGCTCTCAAG
 361 GGCATCGTC GATGACGCT CTCCCTTAG CGACTCCTGC ATTAGGAAGC AGCCCAGTAG
 421 TAGGTTGAGG CCGTTGAGCA CGCCGCCCGC AAGGAATGGT GCATGCAAGG AGATGGCGCC
 481 CAACAGTCCC CGGGCACCGG GGCGCTGCCAC CATAACCCACG CCGAAACAAG CGCTCATGAG
 541 CCCGAAGTGG CGAGCCCGAT CTTCCCCATC GGTGATGTCG GCGATATAGG CGCCAGCAAC
 601 CGCACCTGTG GCGCCGGTGA TGCCGCCAC GATGCGTCCG GCGTAGAGGA TCGAGATCTC
 661 GATCCCGCA AATTAATACG ACTCACTATA GGGAGACCAC AACGGTTTCC CTCTAGATCA
 721 CAAGTTTGTG CAAAAAAAGCT GAACGAGAAA CGTAAAATGA TATAAATATC AATATATTAA
 781 ATTAGATTTT GCATAAAAAA CAGACTACAT AATACTGTAA AACACAAACAT ATCCAGTCAC
 841 TATGGCGGC GCATTAGGCA CCCCAGGCTT TACACTTTAT GTTCCGGCT CGTATAATGT
 901 GTGGATTTTG AGTTAGGATC CGGCGAGATT TTCAGGAGCT AAGGAAGCTA AAATGGAGAA
 961 AAAAATCACT GGATATACCA CGGTGATAT ATCCAATGG CATCGTAAAG AACATTTGA
 1021 GGCATTTCA GTCAGTTGTC AATGTACCTA TAACCAGACC GTTCAGCTGG ATATTACGGC
 1081 CTTTTAAAG ACCGTAAGA AAAATAAGCA CAAGTTTAT CGGGCCTTA TTCACATTCT
 1141 TGCCCGCTG ATGAATGTC ATCCGGAATT CCGTATGGCA ATGAAAGACG GTGAGCTGGT
 1201 GATATGGGAT AGTGTTCACC CTTGTTACAC CGTTTCCAT GAGCAAACIG AAACGTTTTC
 1261 ATCGCTCTGG AGTGAATACC ACGACGATTT CCGGCAGTTT CTACACATAT ATTCGCAAGA
 1321 TGTGGCGTGT TACGGTGAAA ACCTGGCTA TTTCCCTAAA GGGTTTATTG AGAATATGTT
 1381 TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAAGTTT GATTAAACG TGGCCAATAT
 1441 GGACAACCTTC TTGCCCCCG TTTCACCAT GGGCAAATAT TATACGCAAG GCGACAAGGT
 1501 GCTGATGCCG CTGGCGATTC AGGTTCATCA TGCGCTCTGT GATGGCTTCC ATGTCGGCAG
 1561 AATGCTTAAT GAATTACAAC AGTACTGCGA TGAGTGGCAG GGCAGGGCGT AAACGCGTGG
 1621 ATCCGGCTTA CTAAAGGCCA GATAACAGTA TGCGTATTTG CGCGCTGATT TTTGCGGTAT
 1681 AAGAATATAT ACTGATATGT ATACCCGAAG TATGTCAAA AGAGGTGTGC TATGAAGCAG
 1741 CGTATTACAG TGACAGTTGA CAGCGACAGC TATCAGTTGC TCAAGGCATA TATGATGTCA
 1801 ATATCTCCGG TCTGGTAAGC ACAACCATGC AGAATGAAGC CCGTCGTCTG CGTGCAGAAC
 1861 GCTGGAAAGC GGAAAATCAG GAAGGGATGG CTGAGGTCGC CCGGTTTATT GAAATGAACG
 1921 GCTCTTTGC TGACGAGAAC AGGGACTGGT GAAATGCAGT TTAAGGTTA CACCTATAAA
 1981 AGAGAGAGCC GTTATCGTCT GTTTGTGGAT GTACAGAGTG ATATTATTGA CACGCCGGGG
 2041 CGACGGATGG TGATCCCCCT GGCCAGTGCA CGTCCTGCTGT CAGATAAAGT CTCCCGTGAA
 2101 CTTTACCCGG TGGTGCAAT CGGGGATGAA AGCTGGCGCA TGATGACCAAC CGATATGGCC
 2161 AGTGTGCCGG TCTCGTTAT CGGGGAAAGAA GTGGCTGATC TCAGGCCACCG CGAAAATGAC
 2221 ATCAAAAACG CCATTAACCT GATGTTCTGG GGAATATAAA TGTCAGGCTC CCTTATACAC
 2281 AGCCAGTCTG CAGGTCGACC ATAGTGAATG GATATGTTGT GTTTTACAGT ATTATGTTAGT
 2341 CTGTTTTTA TGCAAAATCT AATTAATAT ATTGATATTT ATATCATTTC ACGTTCTCG
 2401 TTCACTTTTC TTGTACAAAG TGGTGATTAT GAGCGATAAA ATTATTACCC TGACTGACGA
 2461 CAGTTTGAC ACGGATGTAC TCAAAGCGGA CGGGCGATC CTGTCGATT TCTGGCAGA
 2521 GTGGTGCCTG CGGTGCAAA TGATGCCCG GATTCTGGAT GAAATCGCTG ACGAATATCA
 2581 GGGCAAACIG ACCGTTGCAA AACTGAACAT CGATCAAAAC CCTGGCACTG CGCCGAAATA
 2641 TGGCATCCGT GGTATCCGA CTCTGCTGCT GTTCAAAAC GGTGAAGTGG CGGCAACCAA
 2701 AGTGGGTGCA CTGTCTAAAG GTCAGTGTAA AGAGTTCTC GACGCTAACCG TGGCCGGTTC
 2761 TGGTTCTGGT GATGACGATG ACAAGGTACG CGGGGATCGA TCCGGCTGCT AACAAAGCCC -

FIGURE 45B

130/240

2821 GAAAGGAAGC TGAGTTGGCT GCTGCCACCG CTGAGCAATA ACTAGCATAA CCCCTTGGGG
 2881 CCTCTAACCG GGTCTTGAGG GGTTTTTGTC TGAAAGGAGG AACTATATCC GGATATCCAC
 2941 AGGACGGGTG TGGTCGCCAT GATCGCTAG TCGATAGTGG CTCCAAGTAG CGAAGGGAGC
 3001 AGGACTGGGC GGCGGCCAAA GCGGTCGGAC AGTGCCTCGA GAACGGGTGC GCATAGAAAT
 3061 TGCATCAACG CATATAGCGC TAGCAGCACG CCATAGTGCAC TGGCGATGCT GTCGGAATGG
 3121 ACGATATCCC GCAAGAGGCC CGGCAGTACCG GGCATAACCA AGCCTATGCC TACAGCATCC
 3181 AGGGTGACGG TGCCGAGGAT GACGATGAGC GCATTGTTAG ATTCATACA CGGTGCCTGA
 3241 CTGCGTTAGC AATTAACTG TGATAAACTA CCGCATTAAA GCTTATCGAT GATAAGCTGT
 3301 CAAACATGAG AATTCTTGAA GACGAAAGGG CCTCGTGATA CGCCTATTTT TATAGGTTAA
 3361 TGTATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT TTTCGGGAA ATGTGCGCGG
 3421 AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA TGAGACAATA
 3481 ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTCCG
 3541 TGTCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCGCTCCTT GTTTTGCTC ACCCAGAAAC
 3601 GCTGGTAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT
 3661 GGATCTAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT
 3721 GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTGTGACG CGGGGCAAGA
 3781 GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAAGTACT CACCAGTCAC
 3841 AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG CCATAACCAC
 3901 GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC
 3961 CGCTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT
 4021 GAATGAAGCC ATACCAAAACG ACGAGCGTGA CACCACGATG CCTGCAGCAA TGGCAACAAAC
 4081 GTTGCACAA CTATTAACGT GCGAAGTACT TACTCTAGCT TCCCGGCAAC AATTAATAGA
 4141 CTGGATGGAG GCGGATAAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG
 4201 GTTTATTGCT GATAAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA TTGCAGCACT
 4261 GGGGCCAGAT GGTAAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGAA GTCAGGCAAC
 4321 TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA AGCATTGGTA
 4381 ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT TTAAAAACTTC ATTTTAATT
 4441 TAAAAGGATC TAGGTGAAGA TCCTTTTGTA TAATCTCATG ACCAAAATCC CTTAACGTGA
 4501 GTTTCTGTT CACTGAGCGT CAGACCCGT AGAAAAGATC AAAGGATCTT CTTGAGATCC
 4561 TTTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAA CCACCGCTAC CAGCGGTGGT
 4621 TTGTTGCCG GATCAAGAGC TACCAACTCT TTTCCGAAG GTAACTGGCT TCAGCAGAGC
 4681 GCAGATACCA AATACTGTCC TTCTAGTGT GCCGTAGTTA GGCCACCACT TCAAGAACTC
 4741 TGTAGCACCG CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG
 4801 CGATAAGTCG TGTCTTACCG GTTGGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG
 4861 GTCGGGCTGA ACGGGGGGTT CGTGCACACA GCCCCAGCTTG GAGCGAACGA CCTACACCGA
 4921 ACTGAGATAC CTACAGCGT AGCTATGAGA AAGCGCCACG CTTCCCGAAG GGAGAAAGGC
 4981 GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG CGCACGAGGG AGCTTCCAGG
 5041 GGGAAACGCC TGGTATCTTT ATAGCTCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG
 5101 ATTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCGACGCA ACGCGGCCCTT
 5161 TTTACGGTTC CTGGCCCTTT GCTGGCCCTT TGCTCACATG TTCTTTCTG CGTTATCCCC
 5221 TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT GATACCGCTC GCGCAGCCG
 5281 AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA GAGCGCCTGA TCGGGTATTT
 5341 TCTCTTACG CATCTGTGCG GTATTTACA CCGCATATAT GGTGCACTCT CAGTACAATC
 5401 TGCTCTGATG CCGCATAGTT AAGCCAGTAT ACACCTCGCT ATCGCTACCGT GACTGGGTCA
 5461 TGGCTCGGCC CGCGACACCCG CCAACACCCG CTGACGCGCC CTGACGGGCT TGTCTGCTCC
 5521 CGGCATCCGC TTACAGACAA GCTGTGACCG TCTCCGGGAG CTGCTATGTGT CAGAGGTTTT
 5581 CACCGTCATC ACCGAAACGC GCGAGGCAGC TGCGGTAAAG CTCATCAGCG TGGTCTGTGAA
 5641 GCGATTCA AATGTCGTC TGTTCATCCG CGTCCAGCTC GTTGAGTTTC TCCAGAACGCG
 5701 TTAATGTCTG GCTTCTGATA AAGCGGGCCA TGTTAAGGGC GGTTTTTCTC TTGTTGGTCA
 5761 CTGATGCCTC CGTGTAAAGGG GGATTTCTGT TCATGGGGGT AATGATACCG ATGAAACGAG
 5821 AGAGGATGCT CACGATACCG GTTACTGATG ATGAAACATGC CCGGTTACTG GAACGTTGTG
 5881 AGGGTAAACA ACTGGCGGTG TGGATGCGGC GGGACCGAGG AAAAATCACT CAGGGTCAAT
 5941 GCCAGCGCTT CGTTAATACA GATGTAGGTG TTCCACAGGG TAGCCAGCAG CATCCCTGCGA
 6001 TGCAGATCCG GAACATAATG GTGCAGGGCG CTGACTTCCG CGTTTCCAGA CTTTACGAAA
 6061 CACGGAAACC GAAGACCAATT CATGTTGTTG CTCAAGGTGCG AGACGTTTTG CAGCAGCAGT
 6121 CGCTTCACGT TCGCTCGCGT ATCGGTGATT CATTCTGCTA ACCAGTAAGG CAACCCCGCC
 6181 AGCCTAGCCG GGTCTCAAC GACAGGAGCA CGATCATGCG CACCCGTGGC CAGGACCCAA
 6241 CGCTGCCGA GATGCGCCGC GTGCGGTGCG TGGAGATGGC GGACGCGATG GATATGTTCT-

F6 U26 45C

131/240

6301 GCCAAGGGTT GGTTTGCAC TTCACAGTTC TCCGCAAGAA TTGATTGGCT CCAATTCTTG
6361 GAGTGGTGAA TCCGTTAGCG AGGTGCCGCC GGCTTCATT CAGGTGAGG TGCCCCGGCT
6421 CCATGCACCG CGACGCAACG CGGGGAGGCA GACAAGGTAT AGGGCGGCAC CTACAATCCA
6481 TGCCAACCCG TTCCATGTGC TCGCCGAGGC GGCATAAATC GCCGTGACGA TCAGCGGTCC
6541 AGTGATCGAA GTAGGCTGG TAAGAGCCGC GAGCGATCCT TGAAGCTGTC CCTGATGGTC
6601 GTCATCTACC TGCCTGGACA GCATGGCCTG CAACGCCGGC ATCCCGATGC CG

FIGURE 45D

132/240

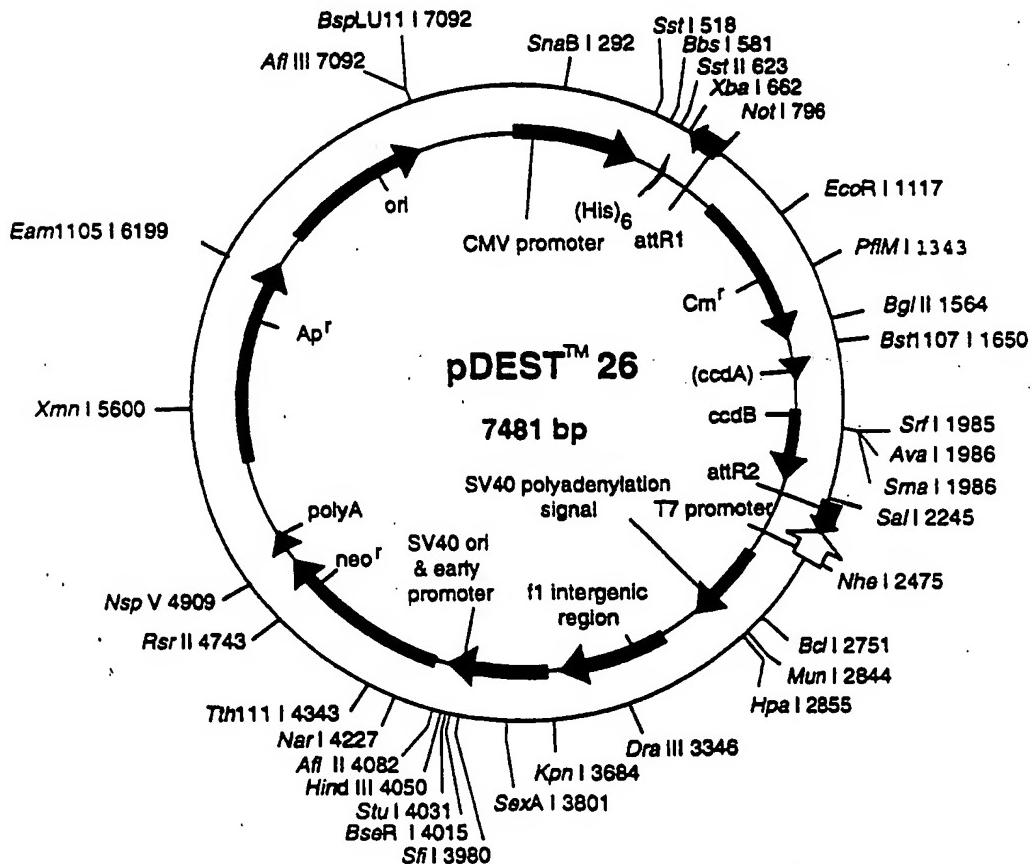
FIGURE 46A

pDEST26 His6 Amino Fusion in pCMV Sport-neo Vector

```

600 ttg acg tca atg gga gtt tgt ttt ggc aee aaa atc aac ggg act ttc caa
     aac tgc agt tac cct caa aca aaa ccc tgg ttt tag tgg ccc tga aag gtt
651 aat gtc gta aca act ccg ccc oat tga cgc aaa tgg gcg gta ggc gtc tac
     tta cag cat tgt tga ggc ggg gta act ggg ttt acc cgc cat ccc cac atg
702 // CMV Promoter → CMV
     ggt ggg agg tct ata taa gca gag ctc gtt tag tga acc gtc aga tgg ttt
     /ccc ccc tcc aca tat att cgt ctc gag caa atc act tgg aag tct ago gga
753 gga gac gcc atc cac gct gtt ttg acc tcc ata gaa gac acc ggg acc gat
     cct ctg cgg tag gtg cga caa aac tgg agg tat ctt ctg tgg ccc tgg cta
804 Start Transl. [A Y Y H H]
     cca gcc tcc gga ctc tag cct agg cgg cgg acc latg gcg tac tac cat nac
     ggt cgg agg cct gag atc gga tcc ggc gcc tgg tac cgc atg atg gta gtg
855 H H H H S R S I S I V K K A 00124/
     cat cac cat cac tct aca tca aca agt ttg tac aaa aaa gct gaa cga gaa
     gta gtg gta gtg aca tct agt tgt tca aac atg ttt ttt cgg ctt gct ctt
     Int ✓

```



133/240

pDEST26 7481 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
492..509	his6
619..519	attR1
752..1411	CmR
1531..1615	inactivated ccdA
1753..2058	ccdB
2099..2223	attR2
2409..2771	SV40 polyA
2966..3421	f1 intergenic region
3485..3903	SV40 promoter
3948..4742	neo
4806..4854	polyA
5265..6125	Apr
6274..6913	ori
7344..385	CMV promoter

1 GTAAACTGCC CACTTGGCAG TACATCAAGT GTATCATATG CCAAGTACGC CCCCTATTGA
 61 CGTCAATGAC GGTAAATGGC CCGCCTGGCA TTATGCCAG TACATGACCT TATGGGACTT
 121 TCCTACTTGG CAGTACATCT ACGTATTAGT CATCGCTATT ACCATGGTGA TGCGGTTTTG
 181 GCAGTACATC AATGGGCGTG GATAGCGTT TGACTCACGG GGATTTCCAA GTCTCCACCC
 241 CATTGACGTC AATGGGAGTT TGTTTGGCA CAAAATCAA CGGGACTTTTC CAAAATGTCG
 301 TAACAACCTCC GCCCCATTGA CGCAAATGGG CGGTAGGCCT GTACGGTGGG AGGTCTATAT
 361 AAGCAGAGCT CGTTTAGTGA ACCGTCAGAT CGCCTGGAGA CGCCATCCAC GCTGTTTG
 421 CCTCCATAGA AGACACCGGG ACCGATCCAG CCTCCGGACT CTAGCCTAGG CCGCGGACCA
 481 TGGCGTACTA CCATCACCAT CACCATCACT CTAGATCAAC AAGTTTGTAC AAAAAAGCTG
 541 AACGAGAAC GTAAAATGAT ATAAATATCA ATATATTAAA TTAGATTTTG CATAAAAAC
 601 AGACTACATA ATACTGTAAA ACACAACATA TCCAGTCACT ATGGCGGCCG CATTAGGCAC
 661 CCCAGGCTTT ACACCTTATG CTTCCGGCTC GTATAATGTG TGGATTTGA GTTAGGATCC
 721 GGCGAGATT TCAGGAGCTA AGGAAGCTAA AATGGAGAAA AAAATCACTG GATATACCAC
 781 CGTTGATATA TCCCAATGGC ATCGTAAAGA ACATTTGAG GCATTTCACT CAGTTGCTCA
 841 ATGTACCTAT AACCAAGACCG TTCAGCTGG A TATTACGGCC TTTTAAAGA CCGTAAAGAA
 901 AAATAAGCAC AAGTTTATC CGGCCTTTAT TCACATTCTT GCCCCGCTGA TGAATGCTCA
 961 TCCGGAATTC CGTATGGCAA TGAAAGACGG TGAGCTGGTG ATATGGATA GTGTTACCC
 1021 TTGTTACACC GTTTCCATG AGCAAATCTGA AACGTTTCA TCGCTCTGG A GTGAATACCA
 1081 CGACGATTC CGGCAGTTTC TACACATATA TTCAGCAAGAT GTGGCGTGT ACGGTAAAAA
 1141 CCTGGCCTAT TCCCCTAAAG GGTTTATG GAATATGTT TTCGTCAG CCAATCCCTG
 1201 GGTGAGTTTC ACCAGTTTG ATTAAACGT GGCCAAATATG GACAACCTCT TCGCCCCCGT
 1261 TTTCACCATG GGCAAATATT ATACGCAAGG CGACAAGGTG CTGATGCCGC TGGCGATTCA
 1321 GGTCATCAT GCGCTCTGTG ATGGCTTCCA TGTCCGGAGA ATGCTTAATG AATTACAACA
 1381 GTACTGCGAT GAGTGGCAGG GCGGGGCGTA AAGATCTGG A TCCGGCTTAC TAAAAGCCAG
 1441 ATAACAGTAT GCGTATTGCG GCGCTGATT TTGCGGTATA AGAATATATA CTGATATGTA
 1501 TACCCGAAGT ATGTCAAAAA GAGGTGTCT ATGAAGCAGC GTATTACAGT GACAGTTGAC
 1561 AGCGACAGCT ATCAGTTGCT CAAGGCATAT ATGATGTCAA TATCTCCGGT CTGGTAAGCA
 1621 CAACCATGCA GAATGAAGCC CGTCGTCTGC GTGCCGAACG CTGGAAAGCG GAAAATCAGG
 1681 AAGGGATGGC TGAGGTGCGC CGGTTTATTG AAATGAACGG CTCTTTGCT GACGAGAACCA
 1741 GGGACTGGTG AAATGCAAGT TAAGGTTAC ACCTATAAAA GAGAGAGCCG TTATCGTCTG
 1801 TTGTTGGATG TACAGAGTGA TATTATTGAC ACGCCACGGC GACGGATGGT GATCCCCCTG
 1861 GCCAGTGCAC GTCTGCTGTC AGATAAAAGTC TCCCCTGAAAC TTTACCCGGT GGTGCATATC
 1921 GGGGATGAAA GCTGGCGCAT GATGACCAAC GATAATGGCA GTGTGCGGGT CTCCGTTATC
 1981 GGGGAAGAAG TGGCTGATCT CAGCCACCGC GAAAATGACA TCAAAACGC CATTAAACCTG
 2041 ATGTTCTGGG GAATATAAAAT GTCAGGCTCC CTTTACACACA GCCAGTCTGC AGGTGGACCA
 2101 TAGTGAATGG ATATGTTG TTTTACAGTA TTATGTAGTC TGTTTTTAT GAAAATCTA
 2161 ATTTAATATA TTGATATTA TATCATTGTA CGTTTCTCGT TCAGCTTCT TGACAAAGT
 2221 GGTTGATCGC GTGCATGCGA CGTCATAGCT CTCTCCCTAT AGTGAGTCGT ATTATAAGCT
 2281 AGGCACGGC CGTCGTTTTA CAACGTCGTG ACTGGAAAAA CTGCTAGCTT GGGATCTTG -

FIGURE 46B

2341 TGAAGGAACC TTACTTCTGT GGTGTGACAT AATTGGACAA ACTACCTACA GAGATTTAAA
 2401 GCTCTAAGGT AAATATAAAA TTTTTAAGTG TATAATGTGT TAAACTAGCT GCATATGCTT
 2461 GCTGCTTGAG AGTTTTGCTT ACTGAGTAG ATTTATGAAA ATATTATACA CAGGAGCTAG
 2521 TGATTCTAAT TGTTTGTGTA TTTTAGATTC ACAGTCCCAA GGCTCATTTG AGGCCCCCTCA
 2581 GTCTCACAG TCTGTTCATG ATCATAATCA GCCATACCAC ATTTGTAGAG GTTTTACTTG
 2641 CTTTAAAAAA CCTCCACAC CTCCCCCTGA ACCTGAAACA TAAAATGAAT GCAATTGTTG
 2701 TTGTTAACCT GTTTATTGCA GCTTATAATG GTTACAAATA AAGCAATAGC ATCACAAATT
 2761 TCACAAATAA AGCATTTCCTT TCACTGCATT CTAGTTGTGG TTTGTCCAAA CTCATCAATG
 2821 TATCTTATCA TGCTGGATC GATCCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC
 2881 GGTTTGCCTA TTGGCTGGCG TAATAGCGAA GAGGCCGCA CCGATCGCCC TTCCCAACAG
 2941 TTGCGCAGCC TGAATGGCGA ATGGGACCG CGCTGTAGCG GCGCAATTAAG CGCGGCGGGT
 3001 GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCCTTC
 3061 GCTTTCTTCC CTCCTTTCT CGCCACGTTG GCCGCTTTG CCCGTCAAGC TCTAAATCGG
 3121 GGGCTCCCTT TAGGGTTCCG ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT
 3181 TAGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTCG CCCTTTGACG
 3241 TTGGAGTCCA CGTTCTTAA TAGTGGACTC TTGTTCCAAA CTGGAACAAC ACTCAACCCCT
 3301 ATCTCGGTCT ATTCTTTGTA TTTATAAGGG ATTTGCGCA TTTCGGCCTA TTGGTTAAAA
 3361 AATGAGCTGA TTTAACAAAT ATTTAACCGG AATTTAACAA AAATATTAAC GTTTACAATT
 3421 TCGCCTGATG CGTATTTC TCCCTACGCA TCTGTGCGGT ATTTCACACC GCATACGCGG
 3481 ATCTGGCAG CACCATGGCC TGAAAATAACC TCTGAAAGAG GAACTTGGTT AGGTACCTTC
 3541 TGAGGGGAA AGAACCGAGCT GTGGAATGTG TGTCAAGTTAG GGTGTGGAAA GTCCCCAGGC
 3601 TCCCCAGCAG GCAGAACGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CAGGTGTGGA
 3661 AAGTCCCCAG GCTCCCCAGC AGGCAGAAGT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA
 3721 ACCATAGTCC CGCCCCCTAAC TCCGCCATC CCGCCCTAA CTCCGCCAG TTCCGCCAT
 3781 TCTCCGCCCC ATGGCTGACT AATTTTTTTT ATTTATGCAG AGGCGCAGGC CGCCTCGGCC
 3841 TCTGAGCTAT TCCAGAAGTA GTGAGGAGGC TTTTTGGAG GCCTAGGCTT TTGCAAAARG
 3901 CTTGATTCTT CTGACACAAAC AGTCTGAAAC TTAAGGCTAG AGCCACCATG ATTGAACAAG
 3961 ATGGATTGCA CGCAGGTTCT CCGGCCGCTT GGGTGGAGAG GCTATTGGC TATGACTGG
 4021 CACAACAGAC AATCGGCTGC TCTGATGCCG CCGTGTCCG GCTGTCAGCG CAGGGCGCC
 4081 CGGTTCTTTT TGTCAAGACC GACCTGTCCG GTGCCCTGAA TGAACTGCAAG GACGAGGCAG
 4141 CGCGGCTATC GTGGCTGGCC ACGACGGGCG TTCCTGCGC AGCTGTGCTC GACGTTGTCA
 4201 CTGAAGCGGG AAGGGACTGG CTGCTATTGG CGCAAGTGCC GGGGCAGGAT CTCTGTCAT
 4261 CTCACCTTGC TCCTGCCGAG AAAGTATCCA TCATGGCTGA TGCAATGCG CGGCTGCATA
 4321 CGCTTGATCC GGCTACCTGC CCATTGACCA ACCAAGCGAA ACATCGCATC GAGCGAGCAC
 4381 GTACTCGGAT GGAAGCCGGT CTGTCGATC AGGATGATCT GGACGAAGAG CATCAGGGC
 4441 TCGGCCAGC CGAACCTTTC GCCAGGCTCA AGGCGCGCAT GCCCCACGGC GAGGATCTCG
 4501 TCGTGACCCA TGGCGATGCG TGCTTGCGA ATATCATGGT GGAAATGGC CGCTTTCTG
 4561 GATTCATCGA CTGTGGCCGG CTGGGTGTGG CGGACCGCTA TCAGGACATA GCGTTGGCTA
 4621 CCCGTGATAT TGCTGAAGAG CTGGCCGGC AATGGGCTGA CCGCTTCCCTC GTGCTTTACG
 4681 GTATCGCCGC TCCCGATTCG CAGCGCATCG CCTTCTATCG CCTTCTTGAC GAGTTCTTCT
 4741 GAGCGGGACT CTGGGGTTGC AAATGACCGA CCAAGCGACG CCCAACCTGC CATCACGATG
 4801 GCCGCAATAA ATAATTTA TTTTCAATTAC ATCTGTGTGT TGGTTTTTG TGTGAATCGA
 4861 TAGCGATAAG GATCCCGTGA TGGTGCACTC TCAGTACAAT CTGCTCTGAT GCCGCATAGT
 4921 TAAGCCAGCC CGCACACCCCG CCAACACCCCG CTGACGCCG CTGACGGGCT TGTCTGCTCC
 4981 CGGCATCCGC TTACAGACAA GCTGTGACCG TCTCCGGGAG CTGCAATGTGT CAGAGGTTTT
 5041 CACCGTCATC ACCGAAACCGC GCGAGACGAA AGGGCCTCGT GATACGCCAA TTTTTATAGG
 5101 TTAATGTCAAT GATAATAATG GTTCTTCTAGA CGTCAGGTGG CACTTTTCCG GGAAATGTGC
 5161 GCGGAACCCC TATTGTTTA TTTTCTAAA TACATTCAA TATGTATCCG CTCATGAGAC
 5221. AATAACCCCTG ATAATGCTT CAATAATATT GAAAAGGAA GAGTATGAGT ATTCAACATT
 5281 TCCGTGTCGC CTTTATTCCC TTTTTGCGG CATTGGCTT TCCTGTTTTT GCTCACCCAG
 5341 AAACGCTGGT GAAAGTAAA GATGCTGAAAG ATCAGTTGGG TGCACGAGTG GGTTACATCG
 5401 AACTGGATCT CAACAGCGGT AAGATCTTG AGAGTTTCG CCCCAGGAA CGTTTCCAA
 5461 TGATGAGCAC TTTAAAGTT CTGCTATGTG GCGCGGTATT ATCCCGTATT GACGCCGGG
 5521 AAGAGCAACT CGGTGCGCCG ATACACTATT CTCAGAATGA CTTGGTTGAG TACTCACCAG
 5581 TCACAGAAAA GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGCACT GCTGCCATAA
 5641 CCATGAGTGA TAACACTGCG GCCAACCTAC TTCTGACAAAC GATCGGAGGA CGGAAGGAGC
 5701 TAACCGCTTT TTTGACAAAC ATGGGGGATC ATGTAACTCG CCTTGATCGT TGGGAACCGG
 5761 AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCAC GATGCCTGTA GCAATGGCAA -

135/240

5821 CAACGTTGCG CAAACTATTA ACTGGCGAAC TACTTACTCT AGCTTCCCGG CAACAATTAA
5881 TAGACTGGAT GGAGGCAGGAT AAAGTTGCAG GACCACCTCT GCGCTCGGCC CTTCCGGCTG
5941 GCTGGTTTAT TGCTGATAAA TCTGGAGCCG GTGAGCGTGG GTCTCGCGT ATCATTGCG
6001 CACTGGGCC AGATGGTAAG CCCTCCCGTA TCGTAGTTAT CTACACGACG GGGAGTCAGG
6061 CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG TGCTCTACTG ATTAAGCATT
6121 GGTAACTGTC AGACCAAGTT TACTCATATA TACTTTAGAT TGATTTAAAAA CTTCATTTTT
6181 AATTTAAAAG GATCTAGGTG AAGATCCTT TTGATAATCT CATGACCAAA ATCCCTTAAC
6241 GTGAGTTTC GTTCCACTGA GCGTCAGACC CCGTAGAAAA GATCAAAGGA TCTTCTTGAG
6301 ATCCTTTTT TCTGCGCGTA ATCTGCTGCT TGCAAACAAA AAAACCACCG CTACCAGCGG
6361 TGGTTTGTGTT GCCGGATCAA GAGCTACCAA CTCTTTTCC GAAGGTAACG GGCTTCAGCA
6421 GAGCGCAGAT ACCAAATACT GTCCCTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA
6481 ACTCTGTAGC ACCGCCTACA TACCTCGCTC TGCTAATCCT GTTACCAAGTG GCTGCTGCCA
6541 GTGGCGATAA GTCGTGTCTT ACCGGGTTGG ACTCAAGACG ATAGTTACCG GATAAGGC
6601 AGCGGTCGGG CTGAACGGGG GGTTCGTGCA CACAGCCCAG CTTGGAGCGA ACGACCTACA
6661 CCGAACTGAG ATACCTACAG CGTGAGCATT GAGAAAGCGC CACGCCCTCCC GAAGGGAGAA
6721 AGGCGGACAG GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCAGC AGGGAGCTTC
6781 CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTCGGGTT TCGCCACCTC TGACTTGAGC
6841 GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG GAAAAAACGCC AGCAACGCC
6901 CCTTTTTACG GTTCCTGGCC TTTTGCTGGC CTTTGCTCA CATGTTCTT CCTGCGTTAT
6961 CCCCTGATTC TGTGGATAAC CGTATTACCG CCTTTGAGTG AGCTGATACC GCTCGCCGCA
7021 GCGGAACGAC CGAGCGCAGC GAGTCAGTGA GCGAGGAAGC GGAAGAGCGC CCAATACGCA
7081 AACCGCCTCT CCCCGCGCGT TGGCCGATTC ATTATGCAG AGCTTGCAAT TCGCGCGTTT
7141 TTCAATATTA TTGAAGCATT TATCAGGGTT ATTGTCAT GAGCGGATAC ATATTTGAAT
7201 GTATTAGAA AAATAAAACAA ATAGGGGTTG CGCGCACATT TCCCCGAAAA GTGCCACCTG
7261 ACGTCTAAGA AACCATTATT ATCATGACAT TAACCTATAA AAATAGGCAGT AGTACGAGGC
7321 CCTTTCACTC ATTAGATGCA TGCGTTACA TAACCTACGG TAAATGGCCC GCCTGGCTGA
7381 CCGCCCAACG ACCCCCGCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT AGTAACGCCA
7441 ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTTAC G

FIGURE 46D

136/240
FIGURE 47A

pDEST 27 GST Amino Fusion in pCMV Sport-neo Vector

CMV Promoter

600 nac ggt ggg agg tct ata taa gca gag ctc gtt tag tga acc gtc aya tcc
 ntg cca ccc tcc aga tat att cgt ctc gag caa atc act tgg dag tct ago

651 cct gga gac gcc atc cac got gtt ttg acc tcc ata gaa gac acc ggg acc
 gga cct ctg cgg tag gtg cga ccc aac tgg agg tat ctt atg tgg ccc tgg

702 gat cca gcc tcc gga ctc tag cct agg cgg acc atg gcc cct ata ata
 cta ggt ogg agg cct gag atc gga tcc gga gcc tgg tcc cgg gga tat gat

753 ggt tat tgg aaa att aag ggc ctt gtg caa ccc act oga ott ctt ttg gaa
 cca ata acc ctt taa ttc ccc gaa cac gtt ggg tga get gaa gaa aac ctt

804 tat ctt gaa gaa aaa tat gaa gag cat ttg tat gag cgg gat gaa gaa ggt gat
 ata gaa ctt ctt ttt ata ctt ctc gta aac ata ctc gcg cta ctt cca cta

— **GST Protein** —

1365 ttt ggt ggt ggc gac cat cct cca aaa tcc gat ctg gtt ccc cgt fct aga
 aaa cca ccc ccg ctg gta gga ggt ttt age cta gac oaa ggo gca aga tot

1416 Eca aca agt ttg tac aaa aaa gct gaa cga gaa acg
 agt ttg tca aac atg ttg ttt cga ctt gct ott tcc

V P R S R

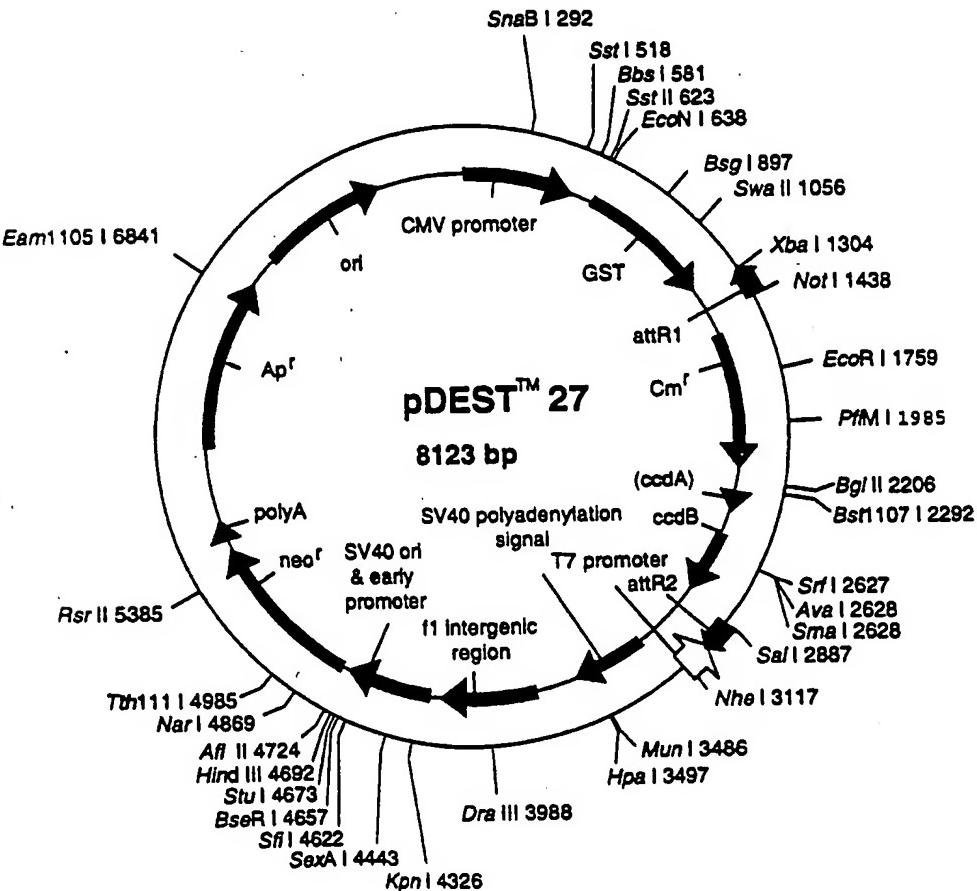
S T S L Y K K A

attR1

Int

M K A A Start

Start Transl/GST



137/240

pDEST27 8123 bp (rotated to position 7800)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
130..793	GST
803..927	attR1
1036..1695	CmR
1815..1899	inactivated ccdA
2037..2342	ccdB
2383..2507	attR2
2693..3055	SV40 polyA
3250..3705	f1 intergenic region
3769..4187	SV40 promoter
4232..5026	neo
5090..5138	polyA
5549..6409	Apr
6558..7197	ori
7628..27	CMV promoter

1 ATAAGCAGAG CTCGTTTAGT GAACCGTCAG ATGCCCTGGA GACGCCATCC ACGCTGTTTT
 61 GACCTCCATA GAAGACACCG GGACCGATCC AGCCTCCGG A CTCTAGCCTA GGCCGCGGAC
 121 CATGGCCCCT ATACTAGGT ATTGGAAAAT TAAGGGCCTT GTGCAACCCA CTCGACTTCT
 181 TTTGGAATAT CTTGAAGAAA AATATGAAGA GCATTTGTAT GAGCGCGATG AAGGTGATAA
 241 ATGGCGAAAC AAAAGTTTG AATTGGGTTT GGAGTTTCCC AATCTTCCTT ATTATATTGA
 301 TGGTGATGTT AAATTAACAC AGTCTATGGC CATCATACGT TATATAGCTG ACAAGCACAA
 361 CATGTTGGGT GGTTGTCCAA AAGAGCGTGC AGAGATTCA ATGCTTGAAG GAGCGGTTTT
 421 GGATATTAGA TACGGTGTTC CGAGAATTGC ATATAGTAAA GACTTTGAAA CTCTCAAAGT
 481 TGATTTCTT AGCAAGCTAC CTGAAATGCT GAAAATGTT GAGATCGTT TATGTCATAA
 541 AACATATTTA AATGGTGATC ATGTAACCCCA TCCTGACTTC ATGTTGTATG ACGCTCTTGA
 601 TGTTGTTTTA TACATGGACC CAATGTGCCT GGATGCGTTT CAAAAATTAG TTTGTTTTAA
 661 AAAACGTATT GAAAGCTATCC CACAAATTGA TAAGTACTTG AAATCCAGCA AGTATATAGC
 721 ATGGCCCTTG CAGGGCTGGC AAGCCACGTT TGGTGGTGGC GACCATCCCTC CAAATCGGA
 781 TCTGGTTCGG CGTTCTAGAT CAACAAGTTT GTACAAAAAA GCTGAACGAG AAACGTTAAA
 841 TGATATAAAT ATCAATATAT TAAATTAGAT TTTGCATAAA AAACAGACTA CATAATACTG
 901 TAAAACACAA CATATCCAGT CACTATGGCG GCCGCATTAG GCACCCCAGG CTTTACACTT
 961 TATGCTTCCG GCTCGTATAA TGTGTGGATT TTGAGTTAGG ATCCGGCGAG ATTTTCAGGA
 1021 GCTAAGGAAG CTAAAATGGA GAAAAAAATC ACTGGATATA CCACCGTTGA TATATCCAA
 1081 TGGCATCGTA AAGAACATT TGAGGCATT CAGTCAGITG CTCATGTAC CTATAACCAG
 1141 ACCGTTTCAGC TGGATATTAC GGCCTTTTA AAGACCGTAA AGAAAAATAA GCACAAGTTT
 1201 TATCCGGCCT TTATTACAT TCTTGGCCGC CTGATGAATG CTCATCCCGA ATTCCGTATG
 1261 GCAATGAAAG ACGGTGAGCT GGTGATATGG GATAGTGTTC ACCCTTGTAA CACCGTTTTC
 1321 CATGAGCAA CTGAAACGTT TTCATCGCTC TGGAGTGAAT ACCACGACGA TTTCCGGCAG
 1381 TTTCTACACA TATATTGCGA AGATGTGGCG TGTACGGTG AAAACCTGGC CTATTTCCCT
 1441 AAAGGGTTTA TTGAGAATAT GTTTTCGTC TCAGCCAATC CCTGGGTGAG TTTCACCAAGT
 1501 TTTGATTTAA ACGTGGCCAA TATGGACAAC TTCTTCGCCC CCGTTTTCAC CATGGGCAA
 1561 TATTATACGC AAGGCACAA GGTGCTGATG CGCCTGGCGA TTCAGGTTCA TCATGCCGTC
 1621 TGTGATGGCT TCCATGTCGG CAGAATGCTT AATGAATTAC AACAGTACTG CGATGAGTGG
 1681 CAGGGCGGGG CGTAAAGATC TGGATCCGGC TTACTAAAAG CCAGATAACA STATCGTAT
 1741 TTGCGCGCTG ATTTTGCGG TATAAGAATA TATACTGATA TGTATACCCG AAGTATGTCA
 1801 AAAAGAGGTG TGCTATGAAG CAGCGTATTA CAGTGCACAGT TGACAGCGAC AGCTATCAGT
 1861 TGCTCAAGGC ATATATGATG TCAATATCTC CGGTCTGGTA AGCACAAACCA TGCAAAATGA
 1921 AGCCCCTCGT CTGCGTCCGG AACGCTGGAA AGCGGAAAT CAGGAAGGGG TGGCTGAGGT
 1981 CGCCCCGGTTT ATTGAAATGA ACGGCTCTTT TGCTGACGAG AACAGGGACT GGTGAAATGC
 2041 AGTTTAAGGT TTACACCTAT AAAAGAGAGA GCGGTTATCG TCTGTTGTG GATGTACAGA
 2101 GTGATATTAT TGACACGCC GGGCACCGA TGGTGTACCC CCTGGCCAGT GCACGTCTGC
 2161 TGTCAGATAA AGTCTCCCGT GAACTTTACC CGGTGGTGCA TATCGGGGAT GAAAGCTGGC
 2221 GCATGATGAC CACCGATATG GCCAGTGTGC CGGTCTCCGT TATCGGGGAA GAAGTGGCTG
 2281 ATCTCAGCCA CGCGAAAAT GACATCAAAA ACGCCATTAA CCTGATGTTC TGGGAATAT-

Figure 47B

138/240

2341 AAATGTCAGG CTCCCTTATA CACAGCCAGT CTGCAGGTG ACCATAGTGA CTGGATATGT
 2401 TGTGTTTAC AGTATTATGT AGTCTGTTT TTATGCAAAA TCTAATTAA TATATTGATA
 2461 TTTATATCAT TTTACGTTTC TCGTTCAGCT TTCTTGTACA AAGTGGTTGA TCGCGTCAT
 2521 GCGACGTCAT AGCTCTCTCC CTATAGTGAG TCGTATTATA AGCTAGGCAC TGGCCGTCGT
 2581 TTTACAACGT CGTGACTGGG AAAACTGCTA GCTTGGGATC TTTGTGAAGG AACCTTACTT
 2641 CTGTGGTGTG ACATAATTGG ACAAACTACC TACAGAGATT TAAAGCTCTA AGGTAAATAT
 2701 AAAATTAAAGTGTATAAT GTGTTAAACT AGCTGCATAT GCTTGCTGCT TGAGAGTTT
 2761 GCTTACTGAG TATGATTAT GAAAATATTA TACACAGGAG CTAGTGATTC TAATTGTTG
 2821 TGTATTAGG ATTCACAGTC CCAAGGCTCA TTTCAGGCC CTCAGTCCTC ACAGTCTGTT
 2881 CATGATCATA ATCAGCCATA CCACATTGT AGAGGTTTTA CTTGCTTTAA AAAACCTCCC
 2941 ACACCTCCCC CTGAACCTGA AACATAAAAAT GAATGCAATT GTTGTGTTA ACTTGTTAT
 3001 TGCAGCTTAT AATGGTTACA AATAAAAGCAA TAGCATCACA AATTTCACAA ATAAAGCATT
 3061 TTTTCACTG CATTCTAGTT GTGGTTGTC CAAACTCATC AATGTATCTT ATCATGTCG
 3121 GATCGATCCT GCATTAATGA ATCGGCCAAC GCGGGGGAG AGGCGGTTTG CGTATTGGCT
 3181 GGCCTAATAG CGAAGAGGCC CGCACCGATC GCCCTTCCA ACAGTTGCGC AGCCTGAATG
 3241 GCGAATGGGA CGCGCCCTGT AGCGGCCGAT TAAGCGCGC GGGTGTGGTG GTTACCGC
 3301 GCGTGACCGC TACACTTGCC AGCGCCCTAG CGCCCGCTCC TTTCGCTTTT TTCCCTTC
 3361 TTCTGCCAC GTTGCAGGCC TTTCCCGTC AAGCTCTAAA TCAGGGGCTC CCTTTAGGGT
 3421 TCCGATTTAG TGCTTTACGG CACCTCGACC CCAAAAAACT TGATTAGGGT GATGGTTCAC
 3481 GTAGTGGGCC ATCGCCCTGA TAGACGGTTT TTTCGCTTT GACGTTGGAG TCCACGTTCT
 3541 TTAATAGTGG ACTCTTGTTG CAAACTGAA CAACACTCAA CCCTATCTCG GTCTATTCTT
 3601 TTGATTATAA AGGGATTGGT CCGATTTCGG CCTATTGGTT AAAAAATGAG CTGATTAAAC
 3661 AAATATTAA CGCGAATTTC AACAATAAT TAACGTTTAC AATTTGCGCT GATGCGGTAT
 3721 TTTCTCCTTA CGCATCTGT CGGTATTTC A CACCGCATAAC GCGGATCTGC GCAGCACCAT
 3781 GGCCTGAAAT AACCTCTGAA AGAGGAACCTT GGTTAGGTAC TTTCGAGGGC GGAAAGAAC
 3841 AGCTGTGGAA TGTGTGTGAG TTAGGGTGTG GAAAGTCCCC AGGCTCCCCA GCAGGCAGAA
 3901 GTATGCAAAG CATGCATCTC ATTAGTCAG CAACCAAGGTG TGGAAAGTCC CCAGGCTCCC
 3961 CAGCAGGCAG AAGTATGCAA AGCATGCATC TCAATTAGTC AGCAACCATA GTCCCGCCCC
 4021 TAACTCCGCC CATCCCGCCC CTAACTCCGC CCAGTTCCGC CCATTCTCCG CCCATGGCT
 4081 GACTAATTAA TTTTATTAT GCAAGGGCG AGGCCGCGCTC GGCTCTGAG CTATTCCAGA
 4141 AGTAGTGAGG AGGCTTTTTT GGAGGCTTAG GCTTTGCAA AAAGCTTGAT TCCTCTGACA
 4201 CAACAGTCTC GAACCTTAAGG CTAGAGCCAC CATGATTGAA CAAGATGGAT TGCACCGCAGG
 4261 TTCTCCGGCC GCTTGGGTGG AGAGGCTATT CGGCTATGAC TGGGCACAAC AGACAATCGG
 4321 CTGCTCTGAT GCCGCGGTGT TCCGGCTGTC AGCGCAGGGG CGCCCGGTTC TTTTTGTC
 4381 GACCGACCTG TCCGGTGCCTC TGAATGAACT GCAGGACGAG GCAGCGCGG TATCGTGGCT
 4441 GGCCACGACG GGGGTTCTT GCGCAGCTGT GCTCGACGTT GTCACTGAAG CGGGAAAGGA
 4501 CTGGCTGCTA TTGGGCGAAG TGCCGGGCA GGATCTCTG TCATCTGAC TTGCTCTGC
 4561 CGAGAAAGTA TCCATCATGG CTGATGCAAT GCGGGGGCTG CATACTGCTT ATCCGGCTAC
 4621 CTGCCCATTC GACCAACAG CGAAACATCG CATCGAGCGA GCACGTACTC GGATGGAAGC
 4681 CGGTCTGTC GATCAGGATG ATCTGGACGA AGAGCATCAG GGGCTCGCGC CAGCCGAAC
 4741 GTTCGCCAGG CTCAGGCAGC GCATGCCGA CGGGGAGGAT CTCGCTGTGA CCCATGGCGA
 4801 TGCCCTGCTTG CCGAATATCA TGGTGGAAA TGGCCGCTT TCTGGATTCA TCGACTGTGG
 4861 CGGGCTGGGT GTGGCGGACC GCTATCAGGA CATAGCGTTG GCTACCCGTG ATATTGCTGA
 4921 AGAGCTTGGC GGCGAATGGG CTGACCGCTT CCTCGTGTAC CGGGCTCCGA CCGCTCCGA
 4981 TTTCGAGCGC ATCGCCCTCT ATCGCCTCT TGACGAGTTT TTCTGAGCGG GACTCTGGGG
 5041 TTTCGAAATGA CCGACCAAGC GACGCCAAC CTGCCATCAC GATGGCCGCA ATAAATATC
 5101 TTTATTAACTA TTACATCTGT GTGTTGGTT TTGTTGTGAA TCGATAGCGA TAAGGATCCG
 5161 CGTATGGTGC ACTCTCAGTA CAATCTGCTC TGATGCCGA TAGTTAAGCC AGCCCGACA
 5221 CCCGCAACA CCCGCTGACG CGCCCTGACG GGCTTGTCTG CTCCCGGCAT CCGCTTACAG
 5281 ACAAGCTGTG ACCGTCTCCG GGAGCTGCAT GTGTCAGAGG TTTTACCGT CATCACCGAA
 5341 ACGCCGAGA CGAAAGGGCC TCGTGATACG CCTATTAA TAGTTAATG TCATGATAAT
 5401 AATGGTTCT TAGACGTCAG GTGGCACCTT CGGGGAAAT GTGCGCGGAA CCCCTATTG
 5461 TTTATTAACTA TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAAT
 5521 GCTTCATAAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA CATTTCGTTG TCGCCCTTAT
 5581 TCCCTTTTTT GCGGCATTTT GCCTTCTGT TTTTGCTCAC CGAGAACGCG TGGTGAAGT
 5641 AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTTAC ATCGAACTGG ATCTCAACAG
 5701 CGGTAAGATC CTTGAGAGTT TTGAGGCCA AGAACGTTT CCAATGATGA GCACTTTAA
 5761 AGTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTG -

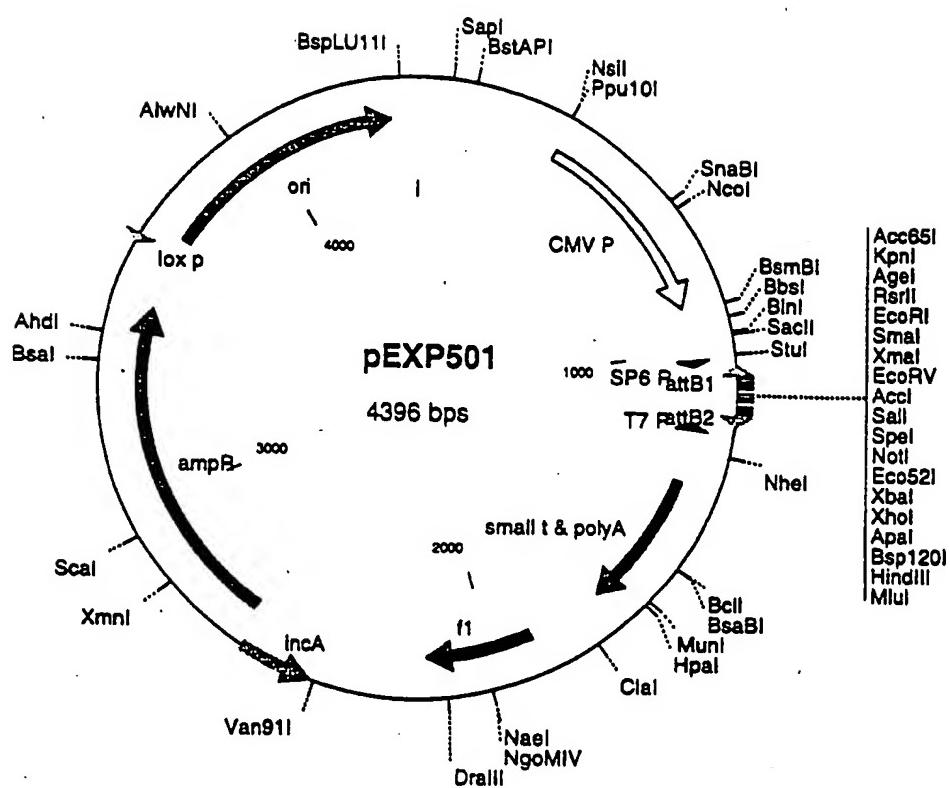
FIGURE 47c

139/240

5821 CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT
 5881 TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC
 5941 TCGGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTGCA
 6001 CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGAA CGGGAGCTGA ATGAAGCCAT
 6061 ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG GCAACAAACGT TGCGAAACT
 6121 ATTAACTGGC GAACTAACCTA CTCTAGCTTC CCGGAAACAA TTAATAGACT GGATGGAGGC
 6181 GGATAAAAGTT GCAGGACCAC TTCTGCCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA
 6241 TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG
 6301 TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG
 6361 AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTGGTAAC TGTCAGACCA
 6421 AGTTTACTCA TATATACTTT AGATTGATTT AAAACTTCAT TTTTAATTAA AAAGGATCTA
 6481 GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCGTTCCA
 6541 CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTCTGCG
 6601 CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA
 6661 TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACAAA
 6721 TACTGTCCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC
 6781 TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG
 6841 TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC
 6901 GGGGGTTCG TGACACACAGC CCAGCTTGA GCGAACGACC TACACCGAAC TGAGATACCT
 6961 ACAGCGTGAN CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAAGGCGG ACAGGTATCC
 7021 GGTAAGCGGC AGGGTGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCGT
 7081 GTATCTTTAT AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG
 7141 CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCCCTTT TACGGTTCC
 7201 GGCCTTTTGC TGGCCTTTG CTCACATGTT CTTTCCCTGCG TTATCCCCTG ATTCTGTGGA
 7261 TAACCGTATT ACCGCCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG
 7321 CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCAATA CGCAAACCGC CTCTCCCCGC
 7381 GCGTTGGCCG ATTCAATTAT GCAGAGCTTG CAATTGCGC GTTTTCAAT ATTATTGAAG
 7441 CATTATTCAG GGTATTGTC TCATGAGCGG ATACATATT GAATGTATT AGAAAAAATAA
 7501 ACAAAATAGGG GTTCCCGCGCA CATTTCCTCCG AAAAGTGCCA CCTGACGTCT AAGAAACCAT
 7561 TATTATCATG ACATTAACCT ATAAAAATAG GCGTAGTACG AGGCCCCCTTC ACTCATTAGA
 7621 TGCAITGCGT TACATAACTT ACGGTAATG GCGGCCCTGG CTGACCGCCC AACGACCCCC
 7681 GCCCATGAC GTCAATAATG ACGTATGTT CCATAGTAAC GCAATAGGG ACTTCCATT
 7741 GACGTCAATG GGTGGAGTAT TTACGGTAAA CTGCCCACTT GGCAGTACAT CAAGTGTATC
 7801 ATATGCCAAG TACGCCCCCT ATTGACGTCA ATGACGGTAA ATGGCCCGCC TGGCATTATG
 7861 CCCAGTACAT GACCTTATGG GACTTTCTA CTTGGCAGTA CATCTACGTA TTAGTCATCG
 7921 CTATTACCAT GGTGATGCGG TTTTGGCAGT ACATCAATGG GCGTGGATAG CGGTTTGA
 7981 CACGGGGATT TCCAAGTCTC CACCCCATTG ACGTCAATGG GAGTTTGTGTT TGGCACCAAA
 8041 ATCAACGGGA CTTTCCAAAA TGTCGTACA ACTCCGCCCC ATTGACGCAA ATGGCCGGTA
 8101 GGC GTGTACG GTGGGAGGTC TAT

FIGURE 47)

Figure 48A: pEXP501: pCMV-SPORT 6 host for attB Libraries



161/240

Figure 4B: pEXP5D1 (cont'd).

Features of the att B cloning vector, pEXP5D1.
 Bases within hatched area are replaced by cDNA in some LTI cDNA libraries.

K68

→ CMV mRNA

---aga gct cgt tta gtg aac cgt cag atc gcc tgg aga cgc cat cca
 ---tct cga gca aat cac ttg gca gtc tag cgg acc tct gcg gta ggt

cgc tgt ttt gac ctc cat aga aga cac cgg gac cga tcc agc ctc
 gcg aca aaa ctg gag gta tct tct gtg gcc ctg gct agg tcg gag

Sst I LTI rev primer

cgg act cta gcc tag gcc ggg gag cgg ata aca att tca cac agg
 gcc tga gat cgg atc cgg cgc ctc gcc tat tgt taa agt gtg tcc

NBTI rev primer

Sph I

SP6 promoter

→ SP6

aaa cag cta tga cca tta ggc cta ttt agg tga cac tat aga aca
 ttt gtc gat act ggt aat cgg gat aaa tcc act gtg ata tct tgt

Int

att B I

Age I Kpn Rsr II

Eco RI

Sma I

agt ttg tac aaa aaa gca ggc agg tgg tat cgg tcc gga att ccc ggg
 tca aac aca ttt ttt cgt ccc act atg gcc agg cct taa ggg ccc

Eco RI Sph I

Spe

Not

Xba

ata / tcg / ccc / agg / agc / tca / tta / gtc / ggc / ggc egc tct aga gta tcc
 tat / age / age / tgc / tcc / agt / gat / dag / ccc / ccc / ggg / aga / tct / cat / agg

Xba

Apa I

Kpn II

Mlu

att B2

Int

ctc gag ggg ccc aag ctt acg cgt acc cag ctt tct tgt aca aag
 gag ccc ccc ggg ttc gaa tgc gaa tgg gtc gaa aga aca tgt ttc

T7

att B2 T7 promoter

Apa I fwd

1272

ttt tac aac gtc gtg act ggg aaa act gct agc ttg gga tct ttg---
 aaa atg tgg cag cac tga ccc ttt tga cga tgg aac cct aga aac---

LTI fwd

142/240

pEXP501 4396 bp

1 CCATTGCCA TTCAGGCTGC GCAACTGTG GGAAGGGCGA TCGGTGCAGGG CCTCTTCGCT
 61 ATTACGCCAG CCAATACGCA AACCGCCTCT CCCCGCGCGT TGCCGATTC ATTAATGCAG
 121 GATCGATCCA GACATGATAA GATACATTGA TGAGTTTGGAA CAAACCACAA CTAGAACATGCA
 181 GTGAAAAAAA TGCTTATTT GTGAAATTG TGATGCTATT GCTTATTTG TAACCATTAT
 241 AAGCTGCAAT AAACAAGTTA ACAACAACAA TTGCATTCA TTTATGTTTC AGGTTCAAGGG
 301 GGAGGTGTGG GAGGTTTTT AAAGCAAGTA AAACCTCTAC AAATGTGGTA TGGCTGATTA
 361 TGATCATGAA CAGACTGTGA GGACTGAGGG GCCTGAAATG AGCCTTGGGA CTGTGAATCT
 421 AAAATACACA AAACAATTAGA ATCACTAGCT CCTGTGTATA ATATTTCTAT AAATCATACT
 481 CAGTAAGCAA AACTCTCAAG CAGCAAGCAT ATGCAGCTAG TTTAACACAT TATACACTTA
 541 AAAATTTAT ATTTACCTTA GAGCTTTAAA TCTCTGTAGG TAGTTTGTC AATTATGTCA
 601 CACCACAGAA GTAAGTTCC TTCACAAAGA TCCCAGCTA GCAGTTTCC CAGTCACGAC
 661 GTTGTAAAAC GACGGCCAGT GCCTAGCTA TAATACGACT CACTATAGGG ACCACTTTGT
 721 ACAAGAAAGC TGGGTACGCG TAAGCTTGGG CCCCTCGAGG GATCCTCTAG AGCGGCCGCG
 781 GACTAGTGAG CTCGTCGACG ATATCCCGGG AATTCCGGAC CGGTACCGC CTGCTTTTT
 841 GTACAAACTT GTTCTATAGT GTCACCTAAA TAGGCCTAAT GGTCACTAGCT GTTCTCTGTG
 901 TGAAATTGTT ATCCGCTCCG CGGCCTAGGC TAGAGTCCGG AGGCTGGATC GGTCCGGTG
 961 TCTTCTATGG AGGTCAAAAC AGCGTGGATG GCGTCTCCAG GCGATCTGAC GGTTCACTAA
 1021 ACGAGCTCTG CTTATATAGA CCTCCCCACCG TACACGCCA CCGCCCCATTT GCGTCAATGG
 1081 GCGGGAGTTG TTACGACATT TTGGAAAGTC CCGTTGATTT TGGTGCAGGG ACAAACTCCC
 1141 ATTGACGTCA ATGGGGTGGG GACTTGGAAA TCCCCGTGAG TCAAACCGCT ATCCACGCC
 1201 ATTGATGTAC TGCCAAAACC GCATCACCAT GGTAATAGCG ATGACTAATA CGTAGATGTA
 1261 CTGCCAAGTA GGAAAAGTCCC ATAAGGTCTAT GTACTGGCA TAATGCCAGG CGGGCCATT
 1321 ACCGTCATTG ACCTCAATAG GGGCGTACT TGGCATATGA TACACTTGAT GTACTGCCA
 1381 GTGGGCAGTT TACCGTAAAT ACTCCACCA TTGACGTCAA TGGAAAGTCC CTATTGGCGT
 1441 TACTATGGGA ACATACGTCA TTATTGACGT CAATGGCGGG GGGTCGTTGG GCGGTCAAGCC
 1501 AGCGGGCCA TTTACCGTAA GTTATGTAAC GACATGCATC TAATGAGTGA AAGGGCCTCG
 1561 TACTACGCCT ATTTTTATAG GTTAATGTCA TGATAATAAT GGTTCTTAG ACGTCAGGTG
 1621 GCACTTTTCG GGGAAATGTG CGCGGAACCC CTATTGTTT ATTTTTCTAA ATACATTCAA
 1681 ATATGTATCC GCTCATGAGA CAATAACCT GATAATGCT TCAATAATAT TGAAAACGC
 1741 GCGAATTGCA AGCTCTGCAT TAATGAATCG GCCAACCGC GGGGAGAGGC GGTTCGCGTA
 1801 TTGGGCCTC TTCCGCTTCC TCGCTACTG ACTCGCTGCG CTCGGTCGTT CGGCTGCC
 1861 GAGCGGTATC AGCTCACTCA AAGGCGTAA TACGTTATC CACAGAATCA GGGGATAACG
 1921 CAGGAAAGAA CATGTGAGCA AAAGGCCAG AAAAGGCCAG GAACCGTAA AAGGCCCGT
 1981 TGCTGGCGTT TTTCATAGG CTCCGCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA
 2041 GTCAGAGGTG GCGAAACCCG ACAGGACTAT AAAGATAACCA GGCGTTTCCC CCTGGAAGCT
 2101 CCCTCGTGCCTCCTCGTCC CGCACCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC
 2161 CTTCGGGAAG CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG
 2221 TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGGCCGAC CGCTGCC
 2281 TATCCGGTAA CTATCGTCTT GAGTCCAACC CGGTAAGACA CGACTTATCG CCACTGGCAG
 2341 CAGCCACTGG TAACAGGATT AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA
 2401 AGTGGTGGCC TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA
 2461 AGCCAGTTAC CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA ACCACCGCTG
 2521 GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG CAGAAAAAAA GGATCTCAAG
 2581 AAGATCCTTT GATCTTTCT ACGGGGTCTG ACGCTCAGTG GAACGAAAAC TCACGTTAAG
 2641 GGATTTGGT CATGCCATAA CTTCGTATAG CATACTTAT ACAGAAGTTAT GGCATGAGAT
 2701 TATCAAAAG GATCTTCACC TAGATCTTT TAAATTAAAA ATGAAGTTT AAATCAATCT
 2761 AAAGTATATA TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA
 2821 TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG ACTCCCCGTC GTGTAGATAA
 2881 CTACGATACG GGAGGGCTTA CCATCTGCC CCAGTGTGTC AATGATACCG CGAGACCCAC
 2941 GCTCACCGGC TCCAGATTTA TCAGCAATA ACCAGGCCAGC CGGAAGGGCC GAGCGCAGAA
 3001 GTGGTCCCTGC AACTTTATCC GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG
 3061 TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC CATTGCTACA GGCATCGTGG
 3121 TGTACGCTC GTCTGGCTT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA TCAAGGGCGAG-

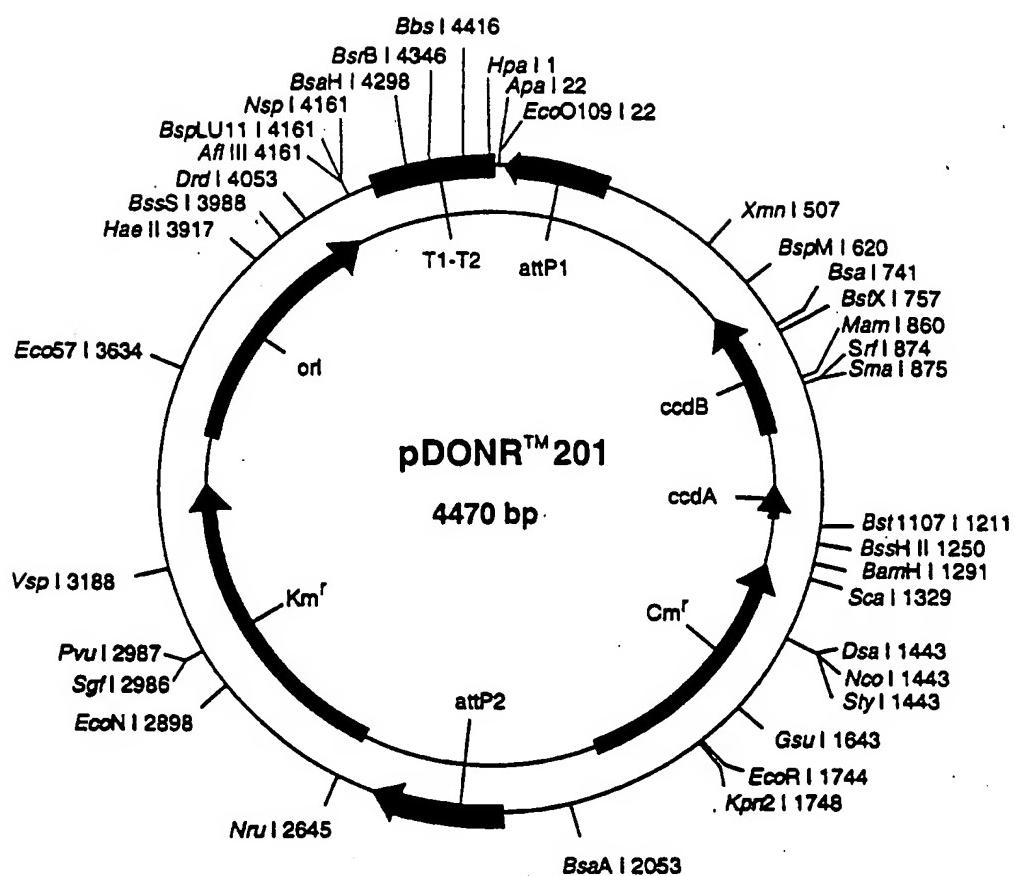
FIGURE 48C

143/240

3181 TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC CTTCGGTCCT CCGATCGTTG
3241 TCAGAAAGTAA GTTGGCCGCA GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC
3301 TTACTGTCAAT GCCATCCGTA AGATGCTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT
3361 TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTGCC CGGCGTCAATA CGGGATAATA
3421 CCGCGCCACA TAGCAGAACT TAAAAAGTGC TCATCATTGG AAAACGTTCT TCAGGGCGAA
3481 AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCACCCA
3541 ACTGATCTTC AGCATCTTT ACTTTCACCA GCGTTCTGG GTGAGCAAAA ACAGGAAGGC
3601 AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC
3661 TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGCCAGGG GTGGGCACAC
3721 ATATTTGATA CCAGCGATCC CTACACAGCA CATAATTCAA TGGCACTTCC CTCTATCGCA
3781 CATCTTAGAC CTTTATTCTC CCTCCAGCAC ACATCGAAC TGCGAGCAA GCCGTTCTCA
3841 CCAGTCCAAG ACCTGGCATG AGCGGATACA TATTTGAATG TATTTAGAAA AATAAACAAA
3901 TAGGGGTTCC GCGCACATT CCCCGAAAAG TGCCACCTGA AATTGTAACAC GTTAATATTT
3961 TGTTAAAATT CGCGTTAAAT TTTTGTAAA TCAGCTCATT TTTTAACCAA TAGGCCAAA
4021 TCGGCAAAAT CCCTTATAAA TCAAAAGAAT AGACCGAGAT AGGGTTGAGT GTTGGTCCAG
4081 TTTGGAACAA GAGTCCACTA TTAAAGAACG TGGACTCCAA CGTCAAAGGG CGAAAAACCG
4141 TCTATCAGGG CGATGGCCCA CTACGTGAAC CATCACCCCTA ATCAAGTTTT TTGGGGTCTGA
4201 GGTGCCGTAA AGCACTAAAT CGGAACCCCTA AAGGGAGCCC CCGATTTAGA GCTTGACGGG
4261 GAAAGCCGGC GAACGTGGCG AGAAAGGAAG GGAAGAAAGC GAAAGGGAGCG GGCGCTAGGG
4321 CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG TAACCACCCAC ACCCGCCGCG CTTAATGCGC
4381 CGCTACAGGG CGCGTC

FIGURE 4B

144/240



pDONR201 4470 bp (rotated to position 3516)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
260..29	attP1
656..961	ccdB
1099..1184	ccdA
1303..1962	CmR
2210..2442	attP2
2565..3374	Kmr
3495..4134	ori

1 GTTAACGCTA GCATGGATCT CGGGCCCCAA ATAATGATT TATTTTGACT GATAUTGACC
 61 TGTCGTTGC AACAAATTGA TGAGCAATGC TTTTTTATAA TGCCAACTTT GTACAAAAAA
 121 GCTGAACGAG AAACGTAAAA TGATATAAAT ATCAATATAT TAAATTAGAT TTTGCATAAA
 181 AAACAGACTA CATAATACTG TAAAACACAA CATATCCAGT CACTATGAAT CAACTACTTA
 241 GATGGTATTA GTGACCTGTG GTCGACCGAC AGCCTTCCAA ATGTTCTTCG GGTGATGCTG
 301 CCAACTTAGT CGACCGACAG CCTTCCAAAT GTTCTTCTCA AACGGAATCG TCGTATCCAG
 361 CCTACTCGCT ATTGTCTCTCA ATGCCGTATT AAATCATAAA AAGAAATAAG AAAAAGAGGT
 421 GCGAGCCTCT TTTTTGTGTG ACAAAATAAA AACATCTACC TATTCTATATA CGCTAGTGTG
 481 ATAGTCCTGA AAATCATCTG CATCAAGAAC AATTCACAA CTCTTATACT TTTCTCTTAC
 541 AAGTCGTTCG GCTTCATCTG GATTTTCAGC CTCTATACCT ACTAAACGTG ATAAAGTTTC
 601 TGTAATTCT ACTGTATCGA CCTGCAGACT GGCTGTGTAT AAGGGAGCCT GACATTATA
 661 TTCCCCAGAA CATCAGGTAA ATGGCGTTTT TGATGTCAATT TTGCGGGTGG CTGAGATCAG
 721 CCACCTCTTC CCCGATAACG GAGACGGGCA CACTGGCCAT ATCGGTGGTC ATCATGCGCC
 781 AGCTTTCATC CCCGATATGC ACCACGGGGT AAAGTTCAAG GGAGACTTTA TCTGACAGCA
 841 GACGTGCACT GGCCAGGGGG ATCACCATCC GTGCCCGGGG CGTGTCAATA ATATCACTCT
 901 GTACATCCAC AAACAGACGA TAACGGCTCT CTCTTTATA GGTGTAAACC TTAAACTGCA
 961 TTTCACCACTG CCCTGTTCTC GTCAAGAAAA GAGCCGTTCA TTCAATAAA CGGGGGGACCC
 1021 TCAGCCATCC CTTCTGTGATT TTCCGCTTTC CAGCGTTCGG CACGCAGACG ACGGGCTTCA
 1081 TTCTGCATGG TTGTGCTTAC CAGACGGGAG ATATTGACAT CATATATGCC TTGAGCAACT
 1141 GATAGCTGTC GCTGTCAACT GTCACTGTAA TACGCTGCTT CATAGCACAC CTCTTTTG
 1201 CATACTTCGG GTATACATAT CAGTATATAT TCTTATACCG CAAAAATCAG CGCGCAAATA
 1261 CGCATACTGT TATCTGGCTT TTAGTAAGCC GGATCCACGC GATTACGCC CGCCCTGCCA
 1321 CTCATCGCAG TACTGTTGTA ATTCAATTAAAG CATTCTGCCG ACATGGAAGC CATCACAGAC
 1381 GGCATGATGA ACCTGAATCG CCAGCGGCAT CAGCACCTTG TCGCCTTGG TATAATATTT
 1441 GCCCATGGTG AAAACGGGGG CGAAGAAGTT GTCCATATTG GCCACGTTA AATCAAAACT
 1501 GGTGAAACTC ACCCAGGGAT TGGCTGAGAC GAAAAACATA TTCTCAATAA ACCCTTTAGG
 1561 GAAATAGGCC AGGTTTTCAC CGTAACACGC CACATCTGC GAATATATGT GTAGAAACTG
 1621 CCGGAAATCG TCGTGGTATT CACTCCAGAG CGATGAAAAC GTTCACTT GCTCATGGAA
 1681 AACGGTGTAA CAAGGGTGA CACTATCCC TATCACCAGC TCACCGTCTT TCATTGCCAT
 1741 ACGGAATTCC GGATGAGCAT TCATCAGGCG GGCAAGAATG TGAATAAAGG CGGGATAAAA
 1801 CTTGTGCTTA TTTTCTTTA CGGTCTTAA AAAGGCCGTA ATATCCAGCT GAACGGCTG
 1861 GTTATAGGTA CATTGAGCAA CTGACTGAAA TGCCCTAAAA TGTTCTTAC GATGCCATTG
 1921 GGATATATCA ACGGTGGTAT ATCCAGTGAT TTTTTCTCC ATTTTAGCTT CCTTAGCTCC
 1981 TGAAAATCTC GATAACTCAA AAAATACGCC CGGTAGTGTAT TTATTTCAT TATGGTGA
 2041 GTTGAACCT CTTACGTGCC GATCAACGTC TCATTTTGC CAAAAGTTGG CCCAGGGCTT
 2101 CCCGGTATCA ACAGGGACAC CAGGATTAT TTATCTGC CGATGATCTT CCGTCACAGG
 2161 TATTATTCTG GCGAAAGTG CGTCGGGTGA TGCTGCCAAC TTAGTCGACT ACAGGGCACT
 2221 AATACCATCT AAGTGTGTA TTCATAGTGA CTGGATATGT TGTGTTTAC AGTATTATGT
 2281 AGTCTGTTT TTATGCAAA TCTAATTAA TATATTGATA TTATATCAT TTTACGTTTC
 2341 TCGTCAGCT TTCTGTACA AAGTGGCAT TATAAGAAAG CATTGCTTAT CAATTGTTG
 2401 CAACGAACAG GTCACTATCA GTCAAATAA AATCATTATT TGCCATCCAG CTGCAGCTCT
 2461 GGCCCGTGTCA TCAAATCTC TGATGTTACA TTGCAACAGA TAAAATATA TCATCATGAA
 2521 CAATAAAACT GTCTGCTTAC ATAAACAGTA ATACAAGGGG TGTTATGAGC CATATTCAAC
 2581 GGGAAACGTC GAGGCCGCGA TTAAATTCCA ACATGGATGC TGATTTATAT GGGTATAAAT
 2641 GGGCTCGCGA TAATGTCGGG CAATCAGGTG CGACAATCTA TCGCTTGAT GGGAAAGCCCG
 2701 ATGCGCCAGA GTTGTCTG AAACATGGCA AAGTAGCGT TGCCAATGAT GTTACAGATG ~

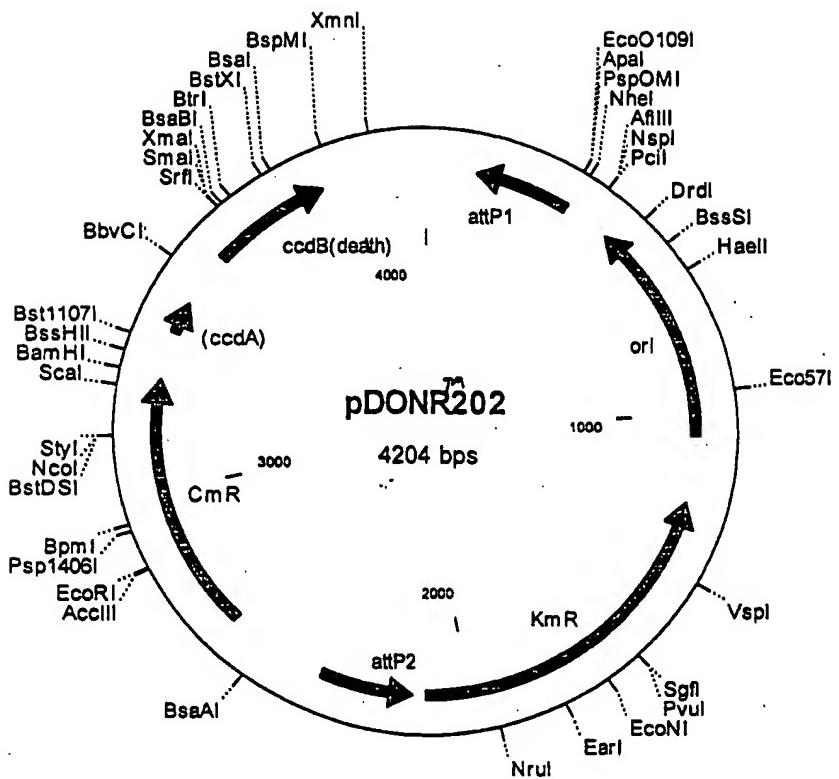
FIGURE 49B

146/240

2761 AGATGGTCAG ACTAAACTGG CTGACGGAAT TTATGCCTCT TCCGACCATC AAGCATTATA
2821 TCCGTACTCC TGATGATGCA TGGTTACTCA CCACTGCGAT CCCCCGGAAAA ACAGCATTCC
2881 AGGTATTAGA AGAATATCCT GATTCAAGTG AAAATATTGT TGATGCGCTG GCAGTGTTC
2941 TGCGCCGGTT GCATTCGATT CCTGTTGTA ATTGTCTTT TAACAGCGAT CGCGTATTC
3001 GTCTCGCTCA GGCAGCAATCA CGAACATGAATA ACGGTTGGT TGATGCGAGT GATTTTGATG
3061 ACGAGCGTAA TGGCTGGCCT GTTGAACAAG TCTGGAAAGA AATGCATAAA CTTTGCCAT
3121 TCTCACCGGA TTCAGTCGTC ACTCATGGTG ATTTCTCACT TGATAACCTT ATTTTGACG
3181 AGGGGAAATT AATAGGTTGT ATTGATGTTG GACCGAGTCGG AATCGCAGAC CGATACCCAGG
3241 ATCTTGCCAT CCTATGGAAC TGCCTCGGT AGTTTCTCC TTCATTACAG AAACGGCTT
3301 TTCAAAAATA TGGTATTGAT AACCTGATA TGAATAAATT GCAGTTTCAT TTGATGCTCG
3361 ATGAGTTTT CTAATCAGAA TTGGTTAATT GGTTGTAACA CTGGCAGAGC ATTACGCTGA
3421 CTTGACGGGA CGGCGCAAGC TCATGACCAA AATCCCTTAA CGTAGTTT CGTTCCACTG
3481 AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTT TTCTGCGCGT
3541 AATCTGCTGC TTGCAAACAA AAAAACCAAC GCTACCAAGCG GTGGTTGTT TGCCGGATCA
3601 AGAGCTACCA ACTCTTTTC CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC
3661 TGTCTTCTA GTGTAGCCGT AGTTAGGCCA CCACCTCAAG AACTCTGTAG CACCGCCTAC
3721 ATACCTCGCT CTGCTAATCC TGTTACCAAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT
3781 TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTGG GCTGAACGGG
3841 GGGTTCGTGC ACACAGCCC GCTTGGAGCG ACGACCTAC ACCGAACCTGA GATACCTACA
3901 GCGTGAGCTA TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT
3961 AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA ACGCCTGGTA
4021 TCTTATAGT CCTGTCGGGT TTGCCCCACT CTGACTTGAG CGTCGATTTT TGTGATGCTC
4081 GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTAC GGTTCCGGC
4141 CTTTGCTGG CCTTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCCTGATT CTGTGGATAA
4201 CCGTATTACC GCTAGCCAGG AAGAGTTGT AGAAACGCAA AAAGGCCATC CGTCAGGATG
4261 GCCTTCTGCT TAGTTTGATG CCTGGCAGTT TATGGCGGGC GTCCCTGCCG CCACCCCTCCG
4321 GGCGTTGCT TCACAAACGTT CAAATCCGCT CCCGGCGGAT TTGTCCTACT CAGGAGAGCG
4381 TTCACCGACA AACAAACAGAT AAAACGAAAG GCCCAGTCTT CCGACTGAGC CTTTCGTTTT
4441 ATTTGATGCC TGGCAGTTCC CTACTCTCGC

FIGURE 49C

147/240
pDONR202 (KanR)
FIGURE 50A:



pDONR202 4204 bp

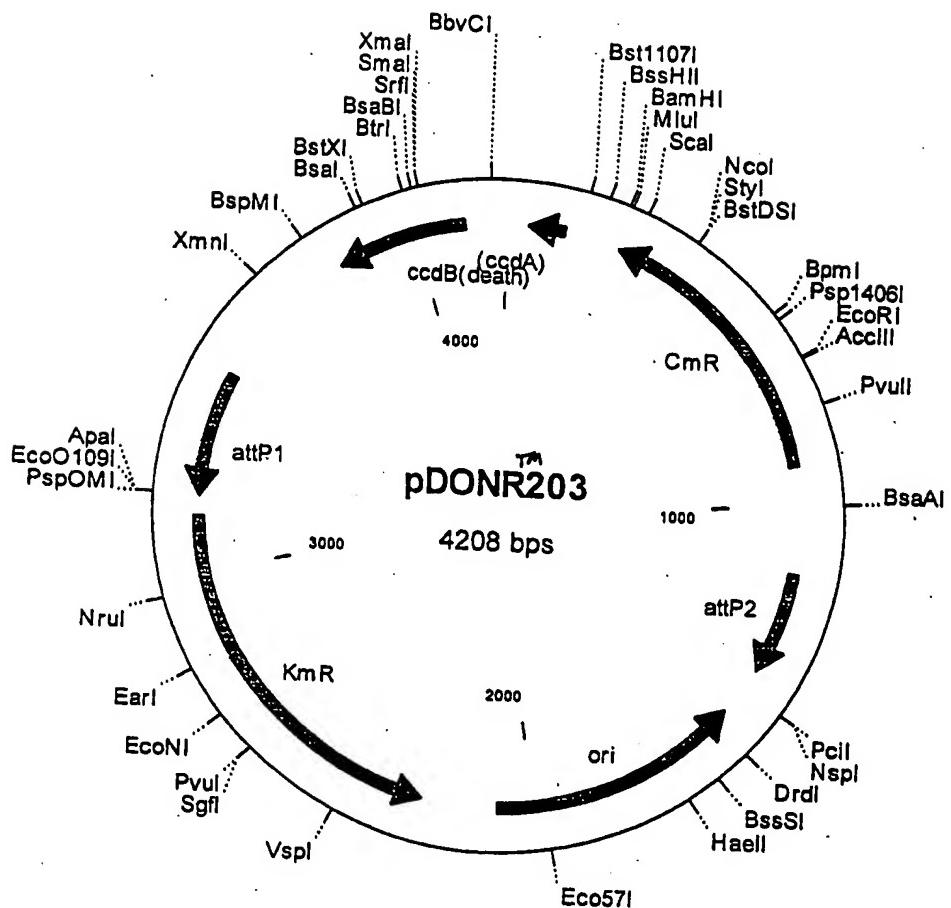
<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
369..127	attP1
486..1059	ori
1228..2107	KmR
2381..2140	attP2
2629..3288	Cmr
3408..3492	inactivated ccdA
3630..3935	ccdB

1 CGGCATTGAG GACAATAGCG AGTAGGCTGG ATACGACGAT TCCGTTTGAG AAGAACATTT
 61 GGAAGGCTGT CGGTCGACTA AGTTGGCAGC ATCACCCGAA GAACATTTGG AAGGCTGTCG
 121 GTCGACTACA GGTCACTAAT ACCATCTAAG TAGTTGATTG ATAGTGACTG GATATGTTGT
 181 GTTTTACAGT ATTATGTTAGT CTGTTTTTA TGCAAAATCT AATTTAATAT ATTGATATTT
 241 ATATCATTTT ACGTTTCTCG TTCACTTTT TGTCACAAAG TTGGCATTAT AAAAAAGCAT
 301 TGCTCATCAA TTTGTTGCAA CGAACAGGTC ACTATCAGTC AAAATAAAAT CATTATTTGG
 361 GGCCCCGAGAT CCATGCTAGC GGTAAATACGG TTATCCACAG AATCAGGGGA TAACGCAGGA
 421 AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG
 481 GCGTTTTTCC ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG
 541 AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC
 601 GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGATAACC TGTCGGCCTT TCTCCCTTCG
 661 GGAAGCGTGG CGCTTTCTCA TAGCTCACCG TGTAGGTATC TCAGTTCGGT GTAGGTCGTT
 721 CGCTCCAAGC TGGGCTGTGT GCACCGAACCC CCCGGTCAGC CCGACCGCTG CGCCTTATCC
 781 GGTAACATATC GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC
 841 ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG
 901 TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA
 961 GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC CGCTGGTAGC
 1021 GGTGGTTTTT TTGTTGCAA GCAGCAGATT ACGCCAGAA AAAAGGATC TCAAGAAGAT
 1081 CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAAACG AAAACTCAGG TTAAGGGATT
 1141 TTGGTCATGA GCTTGCAGCG TCCCGTCAAG TCAGCGTAAT GCTCTGCCAG TGTTACAACC
 1201 AATTAACCAA TTCTGATTAG AAAAACTCAT CGAGCATCAA ATGAAACTGC AATTTTATTCA
 1261 TATCAGGATT ATCAATACCA TATTTTGAA AAAGCCGTTT CTGTAATGAA GGAGAAAAC
 1321 CACCGAGGCA GTTCCATAGG ATGGCAAGAT CCTGGTATCG GTCTGCGATT CCGACTCGTC
 1381 CAACATCAAT ACAACCTATT AATTTCCCCT CGTCACAAAT AGGTTATCA AGTGAGAAAAT
 1441 CACCATGAGT GACGACTGAA TCCGGTGGAA ATGGCAAAAG TTTATGCATT TCTTTCCAGA
 1501 CTTGTTCAAC AGGCCAGCCA TTACGCTCGT CATCAAAATC ACTGCATCA ACCAAACCGT
 1561 TATTCATTG TGATTGCGCC TGAGCGAGAC GAAATACGCG ATCGCTGTTA AAAGGACAAT
 1621 TACAAAACAGG AATCGAATGC AACCAGCGCA GGAACACTGC CAGCGCATCA ACAATATTTT
 1681 CACCTGAATC AGGATATTCT TCTAATACCT GGAATGCTGT TTTTCCGGGG ATCGCAGTGG
 1741 TGAGTAACCA TGCATCATCA GGAGTACGGG TAAAATGCTT GATGGTCGGA AGAGGCATAA
 1801 ATTCCGTCAG CCAGTTTAGT CTGACCATCT CATCTGTAAC ATCATTGGCA ACGCTACCTT
 1861 TGCCATGTTT CAGAAACAC TCTGGCGCAT CGGGCTTCCC ATACAAGCGA TAGATTGTCG
 1921 CACCTGATTG CCCGACATTA TCGCGAGCCC ATTTATACCC ATATAAAATCA GCATCCATGT
 1981 TGGAAATTAA TCGCGGGCTC GACGTTTCCC GTTGAATATG GCTCATAACA CCCCTTGAT
 2041 TACTGTTTAT GTAAGCAGAC AGTTTTATTG TTCATGATGA TATATTTTA TCTTGTGCAA
 2101 TGTAACATCA GAGATTTGA GACACGGGCC AGAGCTGCAG CTGGATGGCA AATAATGATT
 2161 TTATTTGAC TGATAGTGCAC CTGTTGTTG CAACAAATTG ATAAGCAATG CTTTCTTATA
 2221 ATGCCAATTT TGACAAAGAA AGCTGAACGA GAAACGTAAA ATGATATAAA TATCAATATA
 2281 TTAAATTAGA TTTTGCAAA AAAACAGACT ACATAAAACT GTAAAACACA ACATATCCAG
 2341 TCACTATGAA TCAACTACTT AGATGGTATT AGTGCACCTGT AGTCGACTAA GTTGGCAGCA
 2401 TCACCCGACG CACTTTGCGC CGAATAAAATA CCTGTGACGG AAAGTCACCTT CGCAGAAATAA
 2461 ATAAATCCTG GTGTCCCTGT TGATACCGGG AAGCCCTGGG CCAACTTTG GCGAAAATGA
 2521 GACGTTGATC GGCACGTAAG AGGTTCCAAC TTTCACCATA ATGAAATAAG ATCACTACCG
 2581 GGCGTATTTT TTGAGTTATC GAGATTTCA GGAGCTAAGG AAGCTAAAAT GGAGAAAAAA
 2641 ATCACTGGAT ATACCAACCGT TGATATATCC CAATGGCAGTC GTAAAGAACAA TTTTGAGGCA
 2701 TTTCAGTCAG TTGCTCAATG TACCTATAAC CAGACCGTTG AGCTGGATAT TACGGCCTT -

Figure 50B

2761 TTAAAGACCG TAAAGAAAAA TAAGCACAAG TTTTATCCGG CCTTTATTCA CATTCTGCC
2821 CGCCTGATGA ATGCTCATCC GGAATTCGT ATGGCAATGA AAGACGGTGA GCTGGTGATA
2881 TGGGATAGTG TTCACCCCTTG TTACACCGTT TTCCATGAGC AAACGTAAAC GTTTTATCG
2941 CTCTGGAGTG AATACCACGA CGATTTCCGG CAGTTTCTAC ACATATATTC GCAAGATGTG
3001 GCGTGTACG GTGAAAACCT GGCCTATTTC CCTAAAGGGT TTATTGAGAA TATGTTTTC
3061 GTCTCAGCCA ATCCCCTGGGT GAGTTTCAAC AGTTTGATT TAAACGTGGC CAATATGGAC
3121 AACTTCTTCG CCCCCGTTTT CACCATGGGC AAATATTATA CGCAAGGCAGA CAAGGTGCTG
3181 ATGCCGCTGG CGATTCAAGGT TCATCATGCC GTCTGTGATG GTTCCATGT CGGCAGAATG
3241 CTTAACGAAT TACAACAGTA CTGCGATGAG TGGCAGGGCG GGGCGTAATC GCGTGGATCC
3301 GGCTTACTAA AAGCCAGATA ACAGTATGCCG TATTTCGCGC CTGATTTTG CGGTATAAGA
3361 ATATATACGT ATATGTATAC CGAAGTATG TCAAAAAGAG GTGTGCTATG AAGCAGCGTA
3421 TTACAGTGAC AGTTGACAGC GACAGCTATC AGTTGCTCAA GGCATATATG ATGTCATAT
3481 CTCCGGTCTG GTAAGCACAA CCATGCAGAA TGAAGCCCGT CGTCTGCGTG CGAACGCTG
3541 GAAAGCGGAA AATCAGGAAG GGATGGCTGA GGTGCCCCGG TTTATTGAAA TGAACGGCTC
3601 TTTTGTGAC GAGAACAGGG ACTGGTAAA TGCAATTAA GTTTTACACC TATAAAAGAG
3661 AGAGCCGTTA TCGTCTGTTT GTGGATGTAC AGAGTGATAT TATTGACACG CCCGGGCGAC
3721 GGATGGTGTAT CCCCCCTGGCC AGTGCACGTC TGCTGTCAGA TAAAGTCTCC CGTGAACCTT
3781 ACCCGGTGGT GCATATCGGG GATGAAAGCT GGCGCATGAT GACCACCGAT ATGGCCAGTG
3841 TGCCGGTCTC CGTTATCGGG GAAGAAGTGG CTGATCTCAG CCACCGCGAA AATGACATCA
3901 AAAACGCCAT TAACCTGATG TTCTGGGAA TATAAATGTC AGGCTCCCTT ATACACAGCC
3961 AGTCTGCAGG TCGATACAGT AGAAATTACA GAAACTTTAT CACGTTTAGT AAGTATAGAG
4021 GCTAAAATC CAGATGAAGC CGAACGACTT GTAAGAGAAA AGTATAAGAG TTGTGAAATT
4081 GTTCTTGATG CAGATGATTT TCAGGACTAT GACACTAGCG TATATGAATA GGTAGATGTT
4141 TTTATTTGT CACACAAAAA AGAGGCTCGC ACCTTTTTT CTTATTTCTT TTTATGATTT
4201 AATA

FIGURE 50C

FIGURE 5/A pDONR203 (*kanR*)

pDONR203 4208 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
47..131	inactivated ccdA
251..910	CmR
1158..1398	attP2
1509..2082	ori
2251..3130	KmR
3464..3174	attP1
3812..4117	ccdB

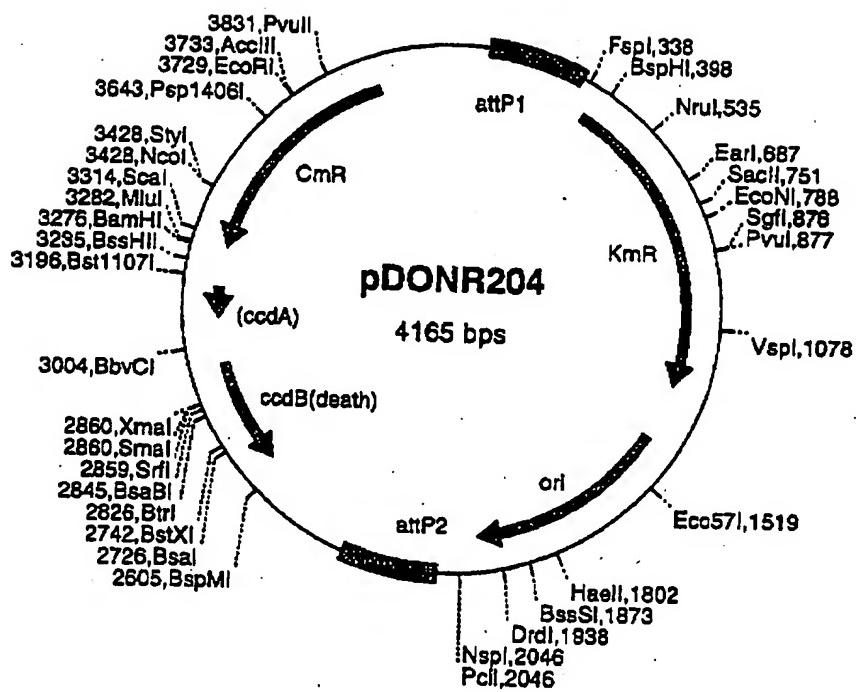
1 GCGTTCCGCA CGCAGACGAC GGGCTTCATT CTGCATGGTT GTGCTTACCA GACCGGAGAT
 61 ATTGACATCA TATATGCCCT GAGCACTGA TAGCTGTCGC TGTCAACTGT CACTGTAATA
 121 CGCTGCTTCA TAGCACACCT CTTTTTGACA TACTTCGGGT ATACATATCA GTATATATTG
 181 TTATACCGCA AAAATCAGCG CGCAAATACG CATACTGTTA TCTGGCTTT AGTAAGCCGG
 241 ATCCACCGGT TTACGCCCCG CCCTGCCACT CATCGCAGTA CTGTTGTAAT TCATTAAGCA
 301 TTCTGCCGAC ATGGAAGCCA TCACAGACGG CATGATGAAC CTGAATCGCC AGCGGCATCA
 361 GCACCTTGTC GCCTTGGTA TAATATTGTC CCATGGTGAAC AACGGGGCG AAGAAGTTGT
 421 CCATATTGGC CACGTTAAA TCAAAACTGG TGAAACTCAC CCAGGGATTG GCTGAGACGA
 481 AAAACATATT CTCATAAAC CTTTAGGAA AATAGGCCAG GTTTTCACCG TAACACGCCA
 541 CATCTGCGA ATATATGTGT AGAAACTGCC GGAAATCGTC GTGGTATTCA CTCCAGAGCG
 601 ATGAAAACGT TTCAAGTTGTC TCATGGAAAA CGGTGTAACA AGGGTGAACA CTATCCATA
 661 TCACCAAGCTC ACCGTCTTTC ATTGCCATAC GGAATTCCGG ATGAGCATTC ATCAGGGGG
 721 CAAGAATGTG AATAAAGGCC GGATAAAACT TGTGTTATT TTTCTTACG GTCTTTAAA
 781 AGGCCGTAAT ATCCAGCTGA ACGGTCTGGT TATAGGTACA TTGAGCAACT GACTGAAATG
 841 CCTCAAAATG TTCTTACGA TGCCATTGGG ATATATCAAC GGTGGTATAT CCAGTGATTT
 901 TTTTCTCCAT TTAGCTTCC TTAGCTCCTG AAAATCTGA TAACCTAAAA AATACGCCCG
 961 GTAGTGATCT TATTCATTA TGGTGAAGT TGGAACCTCT TACGTGCCGA TCAACGTCTC
 1021 ATTTTCGCCA AAAGTTGGCC CAGGGCTTCC CGGTATCAAC AGGGACACCA GGATTTATTT
 1081 ATTCTGCGAA GTGATCTTCC GTCACAGGTA TTTATTCGGC GCAAAGTGCG TCAGGTGATG
 1141 CTGCCAACCT AGTCGACTAC AGGTCACTAA TACCATCTAA GTAGTTGATT CATACTGACT
 1201 GGATATGTTG TGTGTTACAG TATTATGTAG TCTGTTTTT ATGAAAATC TAATTTAATA
 1261 TATTGATATT TATATCATT TACGTTTCTC GTTCAGCTT CTTGTACAAA GTGGCATTAA
 1321 TAAGAAAGCA TTGCTTATCA ATTTGTTGCA ACGAACAGGT CACTATCAGT CAAAATAAAA
 1381 TCATTATTTG CCATCCAGCT AGCGGTAAATA CGGTTATCCA CAGAACATCAGG GGATAACGCA
 1441 GGAAAGAACCA TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAAA GGCCCGCGTTG
 1501 CTGGCGTTTT TCCATAGGCT CCGCCCCCT GACGAGCATC ACAAAAATCG ACGCTCAAGT
 1561 CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC TGGAAAGCTCC
 1621 CTCGTGCGCT CTCTGTTCC GACCCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT
 1681 TCAGGGAAAGCG TGGCGCTTTC TCATAGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC
 1741 GTTCGCTCCA AGCTGGGCTG TGTGCACGAA CCCCCCGTTC AGCCCGACCG CTGCGCCTTA
 1801 TCCGGTAACT ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA
 1861 GCCACTGGTA ACAGGATTAG CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG
 1921 TGGTGGCCTA ACTACGGCTA CACTAGAAGA ACAGTATTTG GTATCTGCGC TCTGCTGAAG
 1981 CCAGTTACCT TCGGAAAAAAG AGTTGGTAGC TCTTGATCCG GCACACAAAC CACCGCTGGT
 2041 AGCGGTGGTT TTTTGTGTTG CAAGCAGCAG ATTACCGCA GAAAAAAAGG ATCTCAAGAA
 2101 GATCCTTCTA TCTTTCTAC GGGGTCTGAC GCTCAGTGGA ACGAAAATC ACCTTAAGGG
 2161 ATTTTGGTCA TGAGCTTGC CGTCCCCTC AAGTCAGCGT AATGCTCTGC CAGTGTAC
 2221 ACCAATTAAC CAATTCTGAT TAGAAAAACT CATCGAGCAT CAAATGAAAC TGCAATTAT
 2281 TCATATCAGG ATTATCAATA CCATATTGTTT GAAAAAGCCG TTTCTGTAAT GAAGGAGAAA
 2341 ACTCACCGAG GCAGTTCCAT AGGATGGCAA GATCCTGGTA TCGGTCTGG ATTCCGACTC
 2401 GTCCAACATC AATACAACCT ATTAATTCC CCTCGTCAAA AATAAGGTTA TCAAGTGAGA
 2461 AATCACCATG AGTGACGACT GAATCCGGTG AGAATGGCAA AAGTTTATGC ATTTCTTCC
 2521 AGACTTGTTC AACAGGCCAG CCATTACGCT CGTCATCAAA ATCACTCGCA TCAACCAAAAC
 2581 CGTTATTCTA TCGTGAATTGC GCCTGAGCGA GACGAAATAC GCGATCGCTG TTAAAGGAC
 2641 AATTACAAAC AGGAATCGAA TGCAACCCGC GCAGGAACAC TGCCAGCGA TCAACAAATAT
 2701 TTTCACCTGA ATCAGGATAT TCTTCTAATA CCTGGAATGC TGTTTTCCG GGGATCGCAG-

2761 TGGTGAGTAA CCATGCATCA TCAGGAGTAC GGATAAAATG CTTGATGGTC GGAAGAGGCA
2821 TAAATTCCGT CAGCCAGTTT AGTCTGACCA TCTCATCTGT AACATCATTG GCAACGCTAC
2881 CTGGCCATG TTTCAGAAC AACTCTGGCG CATCGGGCTT CCCATACAAG CGATAGATTG
2941 TCGCACCTGA TTGCCCAGCA TTATCGCGAG CCCATTATA CCCATATAAA TCAGCATCCA
3001 TGTTGGAATT TAATCGCGC CTCGACGTTT CCCGTTGAAT ATGGCTCATA ACACCCCTTG
3061 TATTACTGTT TATGTAAGCA GACAGTTTTA TTGTTCATGA TGATATATT TTATCTTG
3121 CAATGTAACA TCAGAGATTG TGAGACACGG GCCAGAGCTG CAGCTAGCAT GGATCTGGG
3181 CCCCCAATAA TGATTTTATT TTGACTGATA GTGACCTGTT CGTTGCAACA AATTGATGAG
3241 CAATGCTTTT TTATAATGCC AACTTTGTAC AAAAAAGCTG AACGAGAAAAC GTAAAATGAT
3301 ATAAATATCA ATATATTAAA TTAGATTTC CATAAAAAAC AGACTACATA ATACTGTAAA
3361 ACACAAACATA TCCAGTCACT ATGAATCAC CACTTAGATG GTATTAGTGA CCTGTAGTCG
3421 ACCGACAGCC TTCCAAATGT TCTTCGGGTG ATGCTGCCAA CTTAGTCGAC CGACAGCCTT
3481 CCAAATGTC TTCTCAACG GAATCGCTG ATCCAGCCTA CTCGCTATTG TCCTCAATGC
3541 CGTATTAAT CATAAAAAGA ATAAGAAAA AGAGGTGCGA GCCTCTTTTG TGTGTGACAA
3601 AATAAAAACA TCTACCTATT CATATAGCT AGTGTCAAG TCCTGAAAAT CATCTGCATC
3661 AAGAACAAATT TCACAACTCT TATACTTTT CTTTACAAGT CGTTGGCTT CATCTGGATT
3721 TTCAGCCTCT ATACTTACTA AACGTGATAA AGTTCTGTA ATTTCTACTG TATCGACCTG
3781 CAGACTGGCT GTGTATAAGG GAGCCTGACA TTTATATTCC CCAGAACATC AGGTTAATGG
3841 CGTTTTGAT GTCACTTTCG CGGTGGCTGA GATCAGCCAC TTCTTCCCCG ATAACGGAGA
3901 CGGGCACACT GGGCATATCG GTGGTCATCA TGCGCCAGCT TTCACTCCCCG ATATGCACCA
3961 CGGGGTAAAG TTCACGGGAG ACTTTATCTG ACAGCAGACG TGCACTGGCC AGGGGGATCA
4021 CCATCCGTG CCCGGGCGTG TCAATAATAT CACTCTGTAC ATCCACAAAC AGACGATAAC
4081 GGCTCTCTCT TTATAGGTG TAAACCTAA ACTGCATTTC ACCAGTCCCT GTTCTCGTCA
4141 GCAAAAGAGC CGTTCAATTTC AATAAACCGG GCGACCTCAG CCATCCCTTC CTGATTTC
4201 GCTTTCCA

FIGURE 51C

153/240

Figure 52A POUR204 (KauR)



154/740

pDONR204 4165 bp

1 CGGCATTGAG GACAATAGCG AGTAGGCTGG ATACGACGAT TCCGTTGAG AAGAACATT
 61 GGAAGGCTGT CGGTGACTA CAGGTCACTA ATACCACATCA AGTAGTTGAA TCATAGTGAC
 121 TGGATATGTT GTGTTTACA GTATTATGTA GTCTGTTTT TATGCAAAT CTAATTAAAT
 181 ATATTGATAT TTATATCATT TTACGTTCT CGTTCAGCTT TTTTGACAA AGTTGGCATT
 241 ATAAAAAAAGC ATTGCTTATC AATTTGTTGC AACGAACAGG TCACATATCAG TCAAAATAAA
 301 ATCATTATTT GGGGCCGAG ATCCATGCTA GCTCGAGTGC GCAGGGCCCG TGTCTCAAAA
 361 TCTCTGATGT TACATTGAC AAGATAAAAA TATATCATCA TGAAACAATAA AACTGTCG
 421 TTACATACAA AGTAATACAA GGGGTGTTAT GAGCCATATT CAACGGAAA CGTCTTGCTG
 481 GAGGCCGCGA TTAAATCCA ACATGGATGC TGATTATAT GGGTATAAAT GGGCTCGCGA
 541 TAATGTCGGG CAATCAGGTG CGACAACTTT TCGATTGAT GGGAAAGCCCG ATGCCAGA
 601 GTTGTTCCTG AAACATGGCA AAGGTAGCGT TGCCAATGAT GTTACAGATG AGATGGTCAG
 661 ACTAAACTGG CTGACGGAT TTATGCTCT CGCGACCATC AAGCATTTC TCCGTA
 721 TGATGATGCA TGGTTACTCA CCACATGCGAT CGCGGGAAA ACAGCATTCC AGGTATTAGA
 781 AGAATATCCT GATTCAGGTG AAAATATTGT TGATGCGCTG GCAGTGTCC TGCGCCGGTT
 841 GCATTCGATT CCTGTTGTA ATTGTCCCTT TAACAGCGAT CGCGTATTTC GTCTCGCTCA
 901 GGCGCAATCA CGAACATGAA ACGGTTGGT TGATGCGAGT GATTTGATG ACGAGCGTAA
 961 TGGCTGGCCT GTTGAACAAG TCTGAAAGA AATGCATACG CTTTTGCCAT TCTCACCGGA
 1021 TTCAGTCGTC ACTCATGGTG ATTTCTCACT TGATAACCTT AFTTTTGACG AGGGAAATT
 1081 AATAGGTTGT ATTGATGTTG GACCGAGTCGG AATCGCAGAC CGATAACCAGG ATCTTGCA
 1141 CCTATGGAAC TGCCTCGGTG AGTTTTCTCC TTCAATTACAG AAACGGCTTT TTCAAAATA
 1201 TGGTATTGAT AATCCTGATA TGAATAAAATT GCAGTTTCAT TTGATGCTCG ATGAGTTTT
 1261 CTAATCAGAA TTGGTTAATT GGTTGTAACA CTGGCAGAGC ATTACGCTGA CTTGACGGGA
 1321 CGGCGNCATG ACCAAAATCC CTTAACGTGA GTTTTCGTT CACTGAGCGT CAGACCCCGT
 1381 AGAAAAGATC AAAGGATCTT TTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA
 1441 AACAAAAAAA CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT
 1501 TTTTCCGAAG GTAACGGCT TCAAGCAGAGC GCAGATAACCA AATACTGTCC TTCTAGTGT
 1561 GCGTAGTTA GGCCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT
 1621 AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG GGTTGGACTC
 1681 AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGCACACA
 1741 GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA
 1801 AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGCTGG
 1861 AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT
 1921 CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTTGTGA TGCTCGTCAG GGGGGCGGAG
 1981 CCTATGAAA AACGCCAGCA ACGCGGCCCTT TTTACGGTT CTTGCGCTTT GCTGGCCTTT
 2041 TGCTCACATG TTCTTCCCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCTAG
 2101 CTGGATCGGC AAATAATGAT TTTTATTTGTA CTGATAGTGA CCTGTTCGTT GCAACAAATT
 2161 GATAAGCAAT GCTTTTTAT AATGCCAACT TTGTACAAGA AAGCTGAACG AGAAACGTAA
 2221 AATGATATAA ATATCAATAT ATTAAATTAG ATTTTGACATA AAAAACAGAC TACATAATAC
 2281 TGTAAAACAC AACATATCCA GTCACTATGA TTCAACTACT TAGATGGTAT TAGTGCAC
 2341 TAGTCGACTA AGTTGGCAGC ATCACCCGAC GCACCTTGCG CGAATAAAAT ACCTGTGACG
 2401 GAAGATCACT TCGCAGAATA AATAAACTCT GGTGTCCCTG TTGATACCGG GAAGCCCTGG
 2461 GCCAACTTTT GGCGAAAATG AGACGTTGAT CGGCACATT CACAACCTT ATACTTTCT
 2521 CTTACAAGTC GTTCCGCTTC ATCTGGATTTC TCAGCCTCTA TACTTACTAA ACGTGATA
 2581 GTTTCTGTAA TTTCTACTGT ATCGACCTGC AGACTGGCTG TGTATAACGG AGCCTGACAT
 2641 TTATATTCCC CAGAACATCA GTTTAATGGC GTTTTGATG TCATTTTCGC GGTGGCTGAG
 2701 ATCAGCCACT TCTTCCCCGA TAACGGAGAC CGGCACACTG GCCATATCGG TGGTCATCAT
 2761 GCGCCAGCTT TCATCCCCGA TATGCACCA CGGGTAAAGT TCACGGGAGA CTTTATCTGA
 2821 CAGCAGACGT GCACGGCCA GGGGGATCAC CATCCGTCGC CGGGCGTGT CAATAATATC
 2881 ACTCTGTACA TCCACAAACA GACGATAACG GCTCTCTT TTATAGGTGT AAACCTTAA
 2941 CTGCATTCA CCAGTCCCTG TTCTCGTCAG CAAAAGAGCC GTTCATTCA ATAAACCGGG
 3001 CGACCTCAGC CATCCCTCC TGATTTCCG CTTTCCAGCG TTGGCACGC AGACGACGGG
 3061 CTTCATTCTG CATGTTGTTG CTTACCAAGAC CGGAGATATT GACATCATAT ATGCCTTGAG
 3121 CAACTGATAG CTGTCGCTGT CAACTGTCAC TGTAATACGC TGCTTCATAG CACACCTCTT-

FIGURE 52B

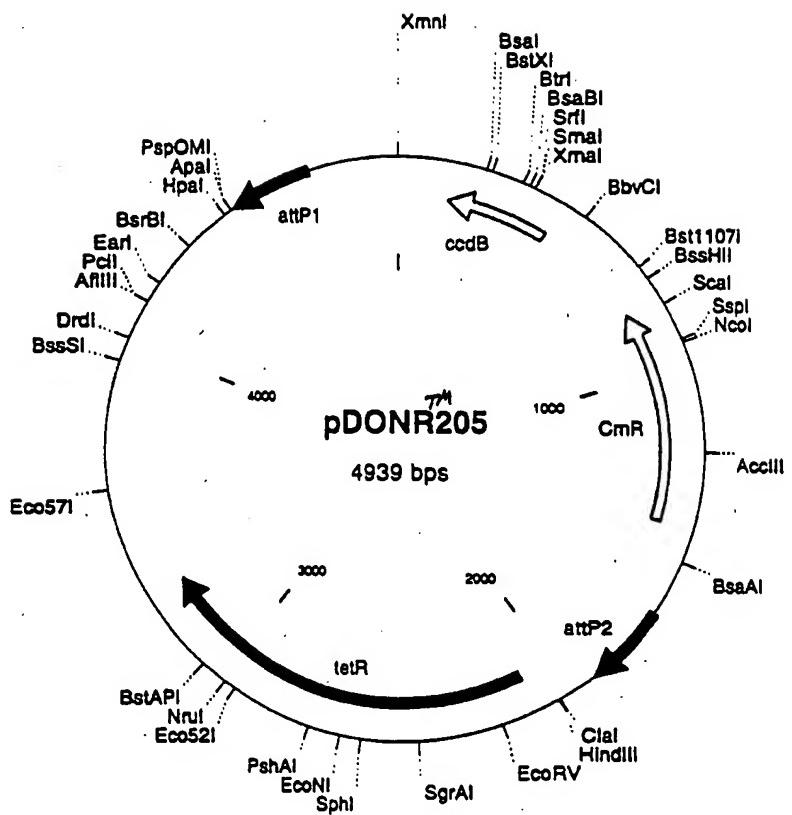
155/240

3181 TTTGACATAC TTCGGGTATA CATATCAGTA TATATTCTTA TACCGCAAAA ATCAGCGCGC
3241 AAATACCGCAT ACTGTTATCT GGCTTTAGT AAGCCGGATC CACCGGTTA CGCCCCGCC
3301 TGCCACTCAT CGCAGTACTG TTGTAATTCA TTAAGCATTC TGCCGACATG GAAGCCATCA
3361 CAGACGGCAT GATGAACCTG AATCGCCAGC GGCATCAGCA CCTTGTGCGCC TTGCGTATAA
3421 TATTTGCCCA TGGTAAAAAC GGGGGCGAAG AAGTTGTCCA TATTGCCAC GTTTAAATCA
3481 AAACCTGGTGA AACTCACCCA GGGATTGGCT GAGACGAAAA ACATATTCTC AATAAACCT
3541 TTAGGGAAAT AGGCCAGGTT TTCAACCGTAA CACGCCACAT CTTGCGAATA TATGTGTAGA
3601 AACTGCCGA AATCGTCGTG GTATTCACTC CAGAGCGATG AAAACGTTTC AGTTTGCTCA
3661 TGGAAAAACGG TGTAACAAGG GTGAAACACTA TCCCATATCA CCAGCTCACC GTCTTCATT
3721 GCCATACCGA ATTCCGGATG AGCAATTCACTC AGGCGGGCAA GAATGTGAAT AAAGGCCGA
3781 TAAAACTTGT GCTTATTTTTT CTTCACGGTC TTTAAAAAGG CCGTAATATC CAGCTGAACG
3841 GTCTGGTTAT AGGTACATTG AGCAACTGAC TGAAATGCCT CAAATGTTTC TTTACGATGC
3901 CATTGGGATA TATCAACGGT GGTATATCCA GTGATTTTTT TCTCCATTTT AGCTTCCTTA
3961 GCTCCTGAAA ATCTCGATAA CTCAAAAAAT ACGCCCGTA GTGATCTTAT TTCAATTATGG
4021 TGAAAAGTTGG AACCTCTTAC TGTTCTTGAT GCAGATGATT TTCAGGACTA TGACACTAGC
4081 ATATATGAAT AGGTAGATGT TTTTATTTTG TCACACAAAA AAGAGGCTCG CACCTCTTTT
4141 TCTTATTTCT TTTTATGATT TAATA

FIGURE 52C

156/240

Figure 53A: pDONR205 (tetR)



(57/240)

pDONR205 4939 bp

GGCATCAGCACCTTGTGCCCTGCGTATAATTTGCCATGGTAAAAACGGGGCGAAG
 AAGTTGTCATATTGCCACGTTAAATCAAACACTGGTAAACTCACCCAGGGATTGGCT
 GAGACGAAAACATATTCTCAATAACCCCTTAGGAAATAGGCCAGGTTTCAACCGTAA
 CACGCCACATCTGCGAATATATGTGTAGAAACTGCGGAAATCGTCGTGGTATTCACTC
 CAGAGCGATGAAAACGTTCAGTTGCTCATGGAAAACGGTAAACAAGGGTGAACACTA
 TCCCATATCACCAGCTACCGTCTTCATTGCCATACGGATTCCGGATGAGCATTCACTC
 AGGGGGCAAGAATGTGAATAAAGCCGATAAAACTTGTGCTTATTTTCTTACGGTC
 TTAAAAAGGCCGTAAATATCCAGCTGAACGGTCTGGTTATAGGTACATTGAGCAACTGAC
 TGAAATGCTCAAAATGTTCTTACGATGCCATTGGATATATCAACGGTGGTATATCCA
 GTGATTTTTCTCCATTAGCTCCTAGCTCCTGAAAATCTGATAACTCAAAAAT
 ACGCCCGGTAGTGTATTTCTTACGGTAAAGTGGAACCTCTACGTGCCGATCA
 ACGTCTCATTTGCCAAAAGTTGCCAGGGCTTCCCGGTATCAACAGGGACACCAGGA
 TTATTTATTCTGCGAAGTGATCTCGTACAGGTATTATTCGGCGAAAGTGCCTG
 GGTGATGCTGCCACTTAGTCGACTACAGGTACTAATACCATCTAAGTAGTTGATTCACT
 AGTGAATGGATATGTTGTTTACAGTATTATGTAGTCTGTTTATGCAAATCTAA
 TTTAATATATTGATATTATATCATTTACGTTCTCGTTCAGCTTCTGTACAAAGTT
 GGCATTATAAGAAAGCATTGCTTATCAATTGGTCAACGAACAGGTACTATCAGTCAA
 AATAAAATCATTATTGCCATCCAGCTGCGCTGGCCCGTGTCTAAAATCTGATG
 TTACATTGACAAGATAAAAATATCATCATCATGAAATTCTCATGTTGACAGCTTATCATC
 GATAAGCTTAATGCCGTAGTTATCACAGTTAAATTGCTAACGCACTGGCAGTCAGGCACCGTGT
 ATGAAATCTAACATGCCGTATCGTCATCCCTGGCACCGTACCCCTGGATGCTGTAGGC
 ATAGGCTGGTTATGCCGTACTGCCGGCTTGTGGGATATGTCATTCCGACAGC
 ATCGCCAGTCACTATGCCGTGCTGCTAGCGTATATGCGTTGATGCAATTCTATGCGCA
 CCCGGTCTCGGAGCACTGTCGACCGCTTGGCGCCCGAGTCCTGCTCGCTCGCTA
 CTGGAGCACTATGACTACGCGATCATGGGACACACCGTCTGTGGATCCTCTAC
 GCCGGACGCGATCGTGGCCGGCATCACCGCGCACAGGTGGGTTGCTGGCGCTATATC
 GCCGACATACCGATGGGAAGATGGGCTGCCACTTGGGCTCATGAGCGCTTGGTTC
 GGCCTGGGTATGGGGCAGGCCCGTGGCGGGGACTGTTGGGCGCATCTCTTGCGAT
 GCACCAATTCTTGCGGGGGCGGTGCTAACGGCTCAACCTACTACTGGCTGCTTCTA
 ATGCAGGAGTCGCTAAAGGGAGAGCTGACCGATGCCATTGAGAGCCTCAACCCAGTC
 AGCTCCTCCGGTGGGGCGGGCATGACTATCGCCTGGCAGCTTGTCTTCTT
 ATCATGCACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCAATTTCGGCGAGGACCGC
 TTGCGTAGCGCAGATGATCGGCTGTGCTTGTGGTATTGGAAATCTGCACTCGC
 CTCGCTAACGCTTGTCACTGGTCCCCACAAACGTTGGCGAGAAGCAGGCCATT
 ATCGCCGGATGGCGGGCGACCGCGTGGCTACGTCCTGCTGGCGTCCGACCGAGGC
 TGGATGGCTTCCCCATTATGATTCTCTCGCTTCCGGCGCATGGGATGCCCGCTTG
 CAGGCCATGCTGTCCAGGCAGGTAGATGACGACCATCAGGGACAGCTCAAGGATCGCTC
 GCGGCTTACCGCTAACGGATTACCAACTTCAAGAATTGGAGCCAATCAATTCTGCGGA
 GAACGTGAATGCGAAACCAACCCCTGGCAGAACATATCCATGCGATGACCAAAATCCC
 TTAACGTGAGTTTCTGTTCCACTGAGCGTCAGACCGCTAGAAAAGATCAAAGGATCTC
 TTGAGATCTTTCTGCGCGTAATCTGCTGCTTGCACAAACAAAAACCCACCGCTACC
 AGCGGTGGTTGTTGCCGATCAAGAGCTACCAACTCTTCCGAAGGTAACGGCTT
 CAGCAGAGCGCAGATAACCAAACCGTCTAGTGAGCGTAGTTAGGCCACCACTT
 CAAGAACTCTGAGACCCGCTACATACCTCGCTCTGCTAATCCTGTTACCGAGGGTGC
 TGCCAGTGGCGATAAGTGTCTTACCGGGTGGACTCAAGACGATAGTTACCGGATAA
 GGCGCAGCGGTGGCTGAACGGGGGGTGTGACACACGCCAGCTGGAGCGAACGAC
 CTACACCGAACCTGAGATAACCTACAGCGTGAGCTATGAGAAAGGCCACGCCAG
 GAGAAAGGCCGACAGGTATCCGTAAGCGCAGGGTGGAAACAGGAGAGCGCACGAGGG
 GCTTCCAGGGGGAAACGCCCTGGTATCTTTATAGTCCTGTCGGGTTTCCGCCACCTCTGACT
 TGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGAAAAACGCCAGCAA-

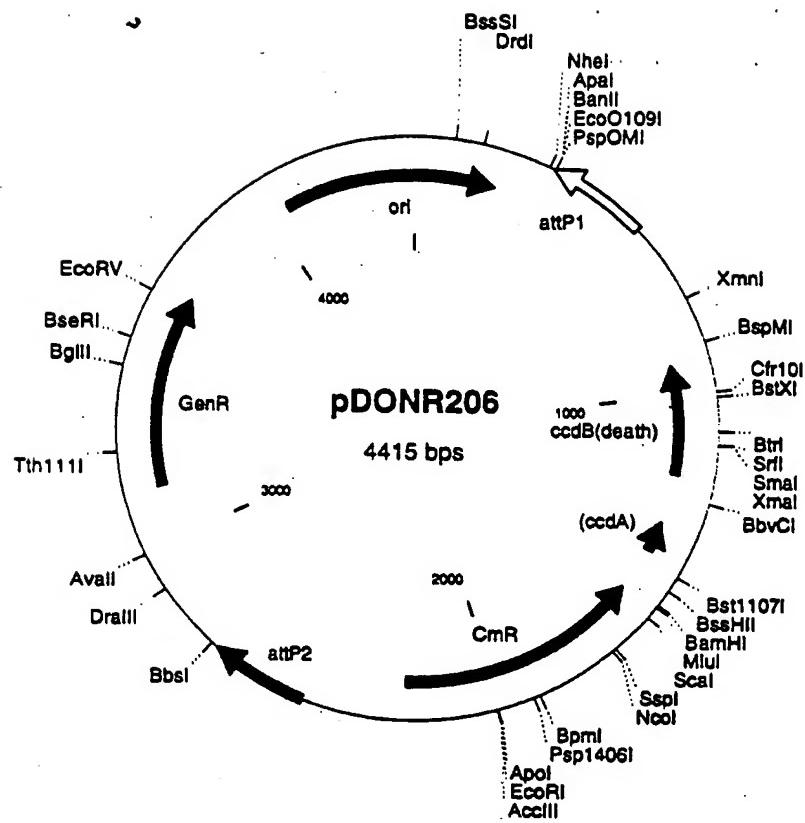
F G UDE 53B

158/260

CGCGGGCTTTTACGGTTCCCTGGCCTTTGCTGGCCTTGCTCACATGTTCTTCCTGC
GTTATCCCCGTATTCTGTGGATAACCGTATTACCGTAGCCAGGAAGAGTTGTAGAAC
GCAAAAAGGCCATCCGTAGGATGGCCTTCTGCTTAGTTGATGCCCTGGCAGTTTATGGC
GGCGTCTGCCCACCCCTCCGGCCGTTGCTCACAACTGTTCAAATCCGCTCCCCGGC
GGATTGTCTACTCAGGAGAGCAGTCACCGACAAACAACAGATAAAAGAAAGGCCAG
TCTTCCGACTGAGCCTTCGTTTATTTGATGCCCTGGCAGTTCCCTACTCTCGCGTTAAC
GCTAGCATGGATCTCGGGCCCCAATAATGATTTATTTGACTGATAGTGACCTGTCG
TTGCAACAAATTGATGAGCAATGCTTTTATAATGCCAACTTTGATCAGAAAAAGCTGAA
CGAGAAACGTAAAATGATATAAATATCAATATATTAAATTAGATTTGATCAGAAAAACAG
ACTACATAAATACTGTAACACACATATCCAGTCACATGAACTACTACTAGATGGT
ATTAGTGACCTGAGTCGACCGACAGCCTCCAAATGTTCTCGGGTGATGCTGCCAACT
TAGTCGACCGACAGCCTCCAATGTTCTCTCAAACGGAATCGTCGTATCCAGCCTACT
CGCTATTGTCCTCAATGCCGTATTAAATCATAAAAAGAAATAAGAAAAAGAGGTGCGAGC
CTCTTTTTGTCGACAAAATAAAACATCTACCTATTCACTACGCTAGTGTATAGTC
CTGAAAATCATCTGCATCAAGAACAAATTCAACACTCTTATACTTTCTTACAAGTCG
TTCGGCTTCATCTGGATTTTCAAGCCTCTATACTACTAAACGTGATAAAGTTCTGTAAT
TTCTACTGTATCGACCTGCAGACTGGCTGTGATAAGGGAGCCTGACATTATATTCCCC
AGAACATCAGGTTAATGGCGTTTGTGTCATTTCGCGGGTGGCTGAGATCAGCCACTT
CTTCCCCGATAACGGAGACCGGCACACTGGCCATATCGGTGGTCATCGGCCAGCTTT
CATCCCCGATATGCACCACCGGGTAAGGTTACGGGAGACTTTATCTGACAGCAGACGTG
CACTGGCCAGGGGGATCACCATCCGTCGCCCGGGCTGTCATAATACTACTCTGTACAT
CCACAAAACAGACGATAACGGCTCTCTTTATAGGTGTAACCTAAACTGCACTTCAC
CAGTCCCTGTTCTCGTCAGCAAAGAGCCGTTCAATTCAATAACCGGGCGACCTCAGCC
ATCCCTTCTGATTTCCGCTTCCAGCGTTGGCACGCAGACGACGGGCTTCATTCTGC
ATGGTTGTGCTTACAGACCGGAGATATTGACATCATATATGCCCTGAGCAACTGATAGC
TGTGCTGTCACTGTCACTGTAATACGCTGCTTCATAGCACACCTTTTGACATACT
TCGGGTATACATATCACTGATATTCTTATACCGAAAAACTAGCGCGAAATACGCGATA
CTGTTATCTGGCTTTAGTAAGCCGGATCCACCGGATTACGCCCGCCCTGCCACTCATC
GCAGTACTGTTGTAATTCACTAACGATTCTGCCGACATGGAAGCCATCACAGACGGCATG
ATGAACTGTAATCGCCAGC

FIGURE 53C

159/240



160/240

pDONR206 4415 bp

CGGCATTGAGGACAATAGCGAGTAGGCTGGATACGACGATTCCGTTGAGAAGAACATTT
 GGAAGGCTGTCGGTCGACTACAGGTCACTAATACCATCTAAGTAGTTGAATCATAGTGAC
 TGGATATGTTGTTTACAGTATTATGTTAGTCGTTTTATGCAAATCTAATTTAA
 ATATTGATATTTATATCATTTCAGTTCTCGTCAGCTTTTGACAAAGTTGGCATT
 ATAAAAAAGATTGCTTATCAATTGTTGCAACGAAACAGGTCACTATCAGTCAAATAAA
 ATCATTATTGGGCCGAGATCCATGCTAGCGTAATACGTTATCCACAGAACAGGG
 GATAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAG
 GCCGCCTGCTGGCGTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGA
 CGCTCAAGTCAGAGGTGGCIAACCCGACAGGACTATAAAAGATACCAGGGCTTCCCCCT
 GGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCGCTTACCGGATACTGTCCGCC
 TTTCTCCCTCGGGAAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTG
 GTGTAGGTGTTGCTCCAAGCTGGGCTGTGTCAGGAACCCCGTTAGCCGGACCCGC
 TGCGCCTTATCCGTAACATATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCA
 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCAGGGTATGTTAGGGCGGTGCTACAGAG
 TTCTTGAAGTGGTGGCTTAACTAGGGCTACACTAGAAGGACAGTATTTGGTATCTGCCT
 CTGCTGAAGCCAGTTACCTCGGAAAAAGAGTGGTAGCTCTTGATCCGGAAACAAACC
 ACCGCTGGTAGCGGTGGTTTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGA
 TCTCAAGAAGATCCTTGATCTTTCTACGGGCTGACGCTCAGTGGAACGAAAATCTA
 CGTTAAGGGATTTGGTCAATGNCGGCTCCGTCAGTCAGCGTAATGCTCTGCCAGTGT
 TACAACCAATTAACCAATTCTGATTAGAAAAAAACTCATCGAGCATCAAATGAAACTGCAAT
 TTATTCAATACAGGATTATCAATACCATATTTTGAAGAACCCGTTCTGTAATGAAGGA
 GAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCCAGTCCG
 ACTCGTCCAAACATCAATACACCTATTAGCCGAGGTCTTCCGATCTCTGAAAGCCAGGGC
 AGATCCGTGACAGCACCTGCGCTAGAAGAACAGCAAGGGCGCCATGCTGACGATGC
 GTGGAGACCGAAACCTTGCCTCGTTCGCCAGGCCAGGACAGAAATGCCGACTTCGCTG
 CTGCCAAGGGTTGCCGGTGACGCAACCCGTGGAAACGGATGAAGGACAGAACCCAGTTG
 ACATAAGCCTGTTGGTGTAAACTGTAATGCAAGTAGCGTATGCCCTGAAAGCCAGGGC
 TCCAGAACCTTGACCGAACGCAAGCGCAGCGGTGTAACGGCGCAGTGGCGTTTTCATGGCTTGT
 TATGACTGTTTTTGACAGTCTATGCCCTGGCATCCAAGCAGCAAGCGCTTACGCC
 GTGGGTCGATGTTGATGTTATGGAGCAGCAACGATGTTACGCGAGCAACGATGTTAC
 GCAGCAGGGCAGTCGCCCTAAACAAAGTTAGGTGGCTCAAGTATGGCATCATTGCCAC
 ATGTTAGGCTCGGCCCTGACCAAGTCAAATCCATGCCCTGCTCTTGATCTTTGCCG
 TGAGTTGGAGACGTTAGCCACCTACTCCAAACATGCCGACTCCGATTACCTCGGGAA
 CTTGCTCCGTAGTAAGACATTCAATGCCCTTGCTGCCCTGCCAACAGAACGGTTGG
 CGCTCTCGCGCTTACGTTCTGCCAGGTTGAGCAGCCGCTAGTGAAGATCTATATCTA
 TGATCTCGCAGTCCTGGCGAGCACCGGAGGCAGGGCATTGCCACCGCGCTCATCAATCT
 CCTCAAGCATGAGGCCAACCGCTGGTCTTATGATCTACGTGCAAGCAGATTACGG
 TGACGATCCCGCAGTGGCTCTCTATACAAAGTTGGGATCACGGGAAGAACGATGACTT
 TGATATGACCCAAGTACGCCACTAACATTGTTCAAGCCGAGATGCCCTCCGGC
 CTAATTCCCTCGTCAAAATAAGGTTATCAAGTGAGAAATCACCAGTGGATGACGACTG
 AATCCGGTGAGAATGCCAAAAGCGTATGCAATTCTTCCAGACTTGTCAACAGGCCAGC
 CATTACGCTCGTCAAAATCACTCGCATCAACAAACCGTTATTGTTGATGCG
 CCTGAGCGAGCAGAAATACCGGATCGCTGTTAAAGGACAAATTACAAACAGGAATCGAAT
 GCAACCGGGCGAGGAAACACTGCCAGCGCATCAACAAATTCTTCCAGCTGAGTGGATATT
 CTCTCAATACCTGGATGCTTTCCCGGGATCGCAGTGGTGAGTACCATGCGATCAT
 CAGGAGTACGGATAAAATGCTGATGGTGGAGGCAACGCTAAATTCCGTCAGCCAGTTA
 GTCTGACCATCTCATCTGTAACATCATTGGCAAGCTACCTTGGCATGTTGAGAACAA
 ACTCTGGCGCATGGGCTCCCATACAATCGAAAGATTGTCGACCTGATTGCCCAGAT
 TATCGCGAGCCCATTATACCCATATAAAATCGCATCCATGTTGGAAATTAAATCGGGCC
 TCCAGCAAGACGTTCCGTTGAATATGGCTCATAACACCCCTGTATTACTGTTATGT
 AAGCAGACAGTTTATGTTGATGATGATATAATTCTTGCAATGAACTGATGAGTTATTTG
 GATTTGAGACACGGGCCNGCGCACTGCAGCTGGATCGGAAATAATGATTATTTG
 ACTGATAGTGCACCTGTTGCAACAAATTGATAAGCAATGCTTTTATAATGCCAAC -

FIGURE 54B

(61 / 240)

TTTGTACAAGAAAGCTGAACGAGAACGTAATGATATAAATATCAATATATTAAATTA
GATTTGCATAAAAACAGACTACATAATACTGTAACACAAACATATCCAGTCAGTATG
ATTCAACTACTTAGATGGTATTAGTGACCTGTAGTCAGTAAGTTGGCAGCATACCGA
CGCACCTTGCGCCGAATAAAATACCTGTGACGGAAGATCACTTTCGAGAATAAAATACCC
TGGTGTCCCTGTTGATACCGGGAAAGCCCTGGGCAACTTTGGCAGAAATGAGACGTTGA
TCGGCACGTAAGAGGTTCCAACCTTCACCAATAATGAAATAAGATCACTACCGGGCGTATT
TTTGAGTTATCGAGATTTCAAGGAGCTAAGGAAGCTAAATGGAGAAAAAAATCACTGG
ATATACCAACCGTTGATATATCCATGGCATCGTAAAGAACATTTGAGGCATTCAGTC
AGTTGCTCAATGTACCTATAACCAGACCGTTCAAGCTGGATATTACGGCTTTAAAGAC
CGTAAAGAAAAATAAGCACAAAGTTTATCCGGCTTATTACACATTCTGGGCTTGAT
GAATGCTCATCCGAATTCCGTATGGCAATGAAAGACGGTAGCTGGTGTATGGGATAG
TGTTCACCTTGTACACCGTTCCATGAGCAAACCTGAAACGTTTCACTGCTCTGGAG
TGAATACCAACGACGATTCCGGCAGTTCTACACATATAATCGCAAGATGTGGCGTGT
CGGTAAAACCTGGCTATTCCTAAAGGGTTATTGAGAATATGTTTTCGTCTCAGC
CAATCCCTGGGTGAGTTACACAGTTTGATTTAAACGTGGCAATATGGACAACCTCTT
CGCCCCCGTTTCACCATGGCAAATAATTACGCAAGGCGACAAGGTGCTGATGCCGCT
GGCGATTCAAGGTCATCATGGCGCTGTGATGGCTTCCATGCGCAGAAATGCTTAATGA
ATTACAACAGTACTGCGATGAGTGGCAGGGCGGGCGTAAACGCGTGGATCCGGCTACT
AAAAGCCAGATAACAGTATGCGTATTTGCGCGCTGATTTTGGGTATAAGAATATATAC
TGATATGTATACCGAAGTATGTCAAAAAGAGGTGCTATGAAAGCAGCGTATTACAGTG
ACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAATATCTCCGGTC
TGGTAAGCACAACCATGCGAATGAAGCCCGTCGTCTGCGTGCGAACGCTGGAAAGCGG
AAAATCAGGAAGGGATGGCTGAGGTGCCCCGGTTATTGAAATGAACGGCTTTTGTG
ACGAGAAACAGGGACTGGTGAATGCAAGTTAAGGTTACACCTATAAAAGAGAGAGCGT
TATCGTCTGTTGTGGATGTACAGGTGATATTATTGACACGCCCGGGCGACGGATGGT
ATCCCCCTGGCCAGTGCACGTCTGCTGTCAGATAAAAGTCTCCCGTGAACCTTACCCGGT
GTGCATATCGGGGATGAAAGCTGGCGCATGATGACCAACCGATATGGCCAGTGTGCCGGTC
TCCGTTATCGGGGAAGAAGTGGCTGATCTCAGCCACCGCGAAAATGACATCAAAACGCC
ATTAACCTGATGTTCTGGGAATATAAAATGTCAGGCTCCGGTATACACAGCCAGTCTGCA
GGTCGATACAGTAGAAATTACAGAAACTTTACACGTTAGTAAGTATAGGGCTGAAAAA
TCCAGATGAAGCCGAACGACTTGTAAAGAGAAAAGTATAAGAGTGTGAAATTGTTCTTGA
TGCAGATGATTTCAAGGACTATGACACTAGCATATATGAATAGGTAGATGTTTATTT
GTCACACAAAAAGAGGCTCGCACCTTTTCTTATTCTTTATGATTTAATA

FIGURE 54 C

162/240

Figure 55 An Entry (pENTR7) Clone of CAT Subcloned into pDEST2

1021 cgg ata aca att tca cac agg aaa cag acc atg tgg tac ttc cat cac cat
 gcc tat tgt taa agt gtg tcc ttt gtc egg tac age atg atg gta gtg gta

→ Start translation

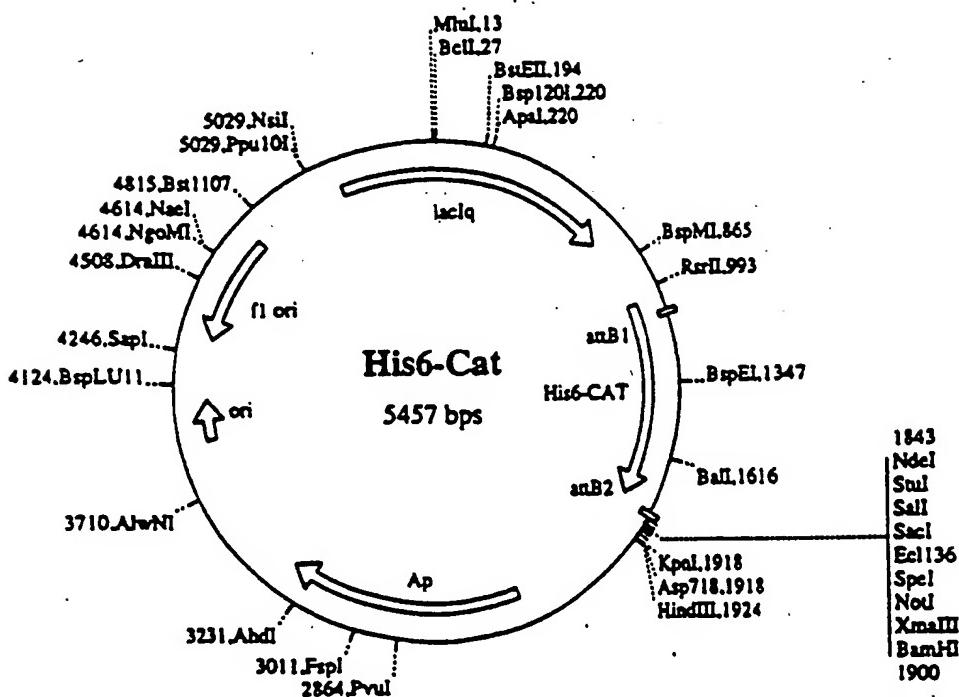
1072 His6 ← His His His Gly Ile Thr Ser Leu Tyr Lys Lys Ala Glu Phe Glu Asn Leu
 cac cat cac cac ggc acc tca agt tgg tac aaa aaa gca ggc ttt gaa aac ctg
 gtg gta gtg ccg tag tgc tea aac atg EEE ttt cgt ccg aha ctt ttg gac

attB1

From pDEST2 From pENTR7

TEV protease → Start CAT

1123 Tyr Phe Gln ↓ Gly Thr Met Gly Lys Lys Ile Thr Glu Tyr Thr Val Asp
 tat ccc caa gga acc atg gag aca aca acc act gga tat acc acc gtt gat
 ata aaa gtt cct tgg tac ctc ttt ttt tag tga ccc ata tgg tgg caa cta



163/240

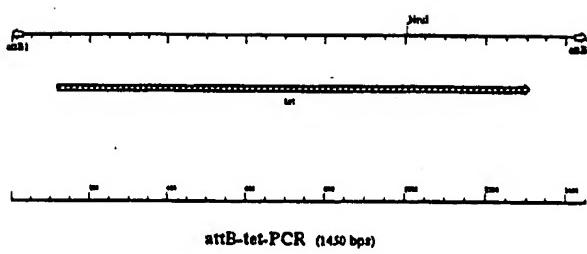
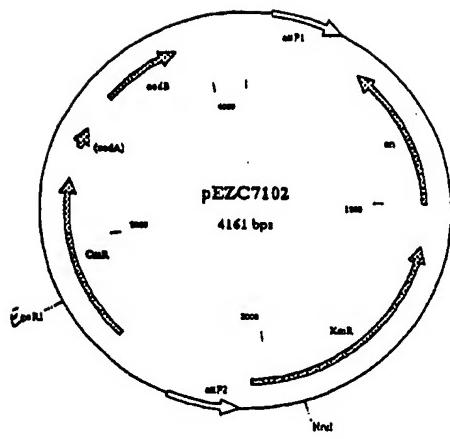


FIGURE 56

164/240

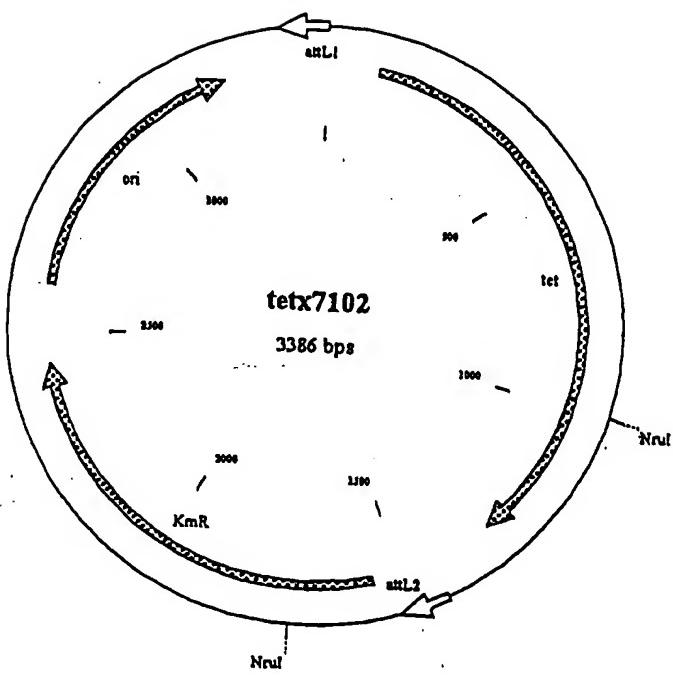


FIGURE 57

165/240

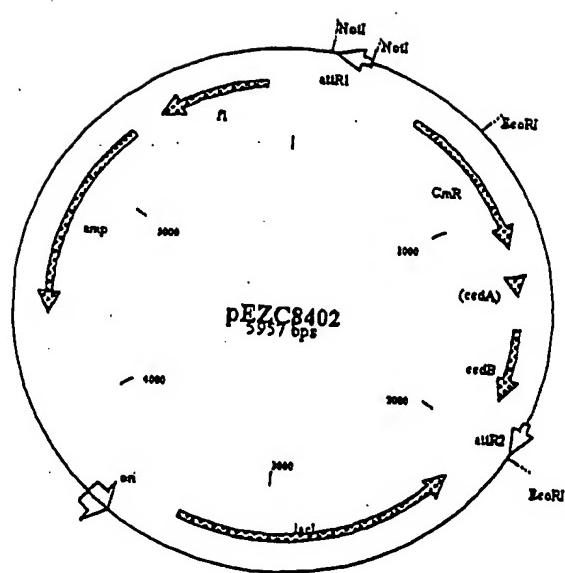


FIGURE 58

166/260

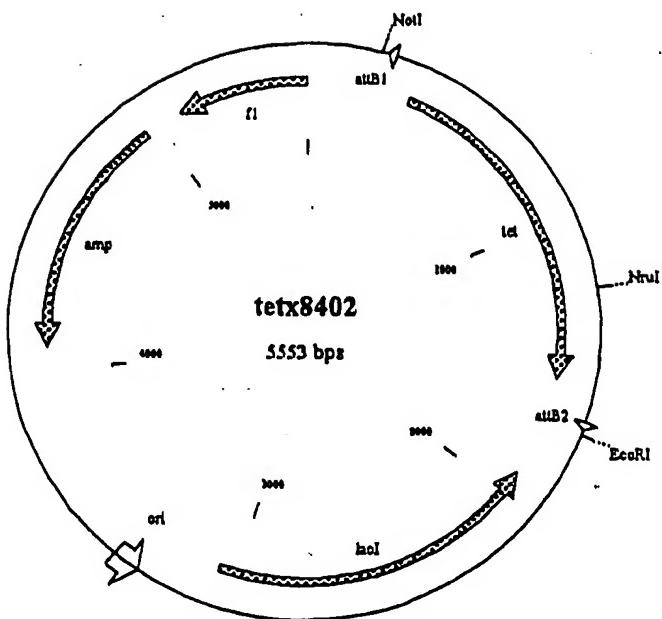


FIGURE 59

167/240

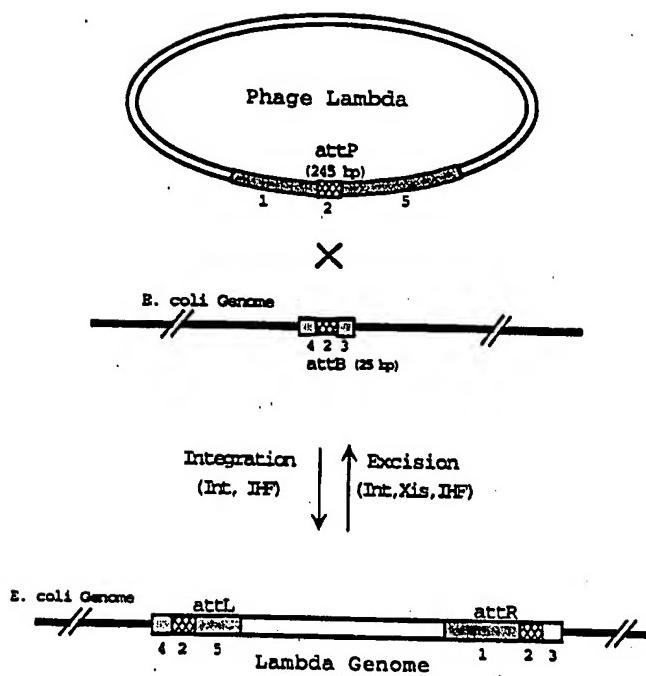


FIGURE 60

168/240

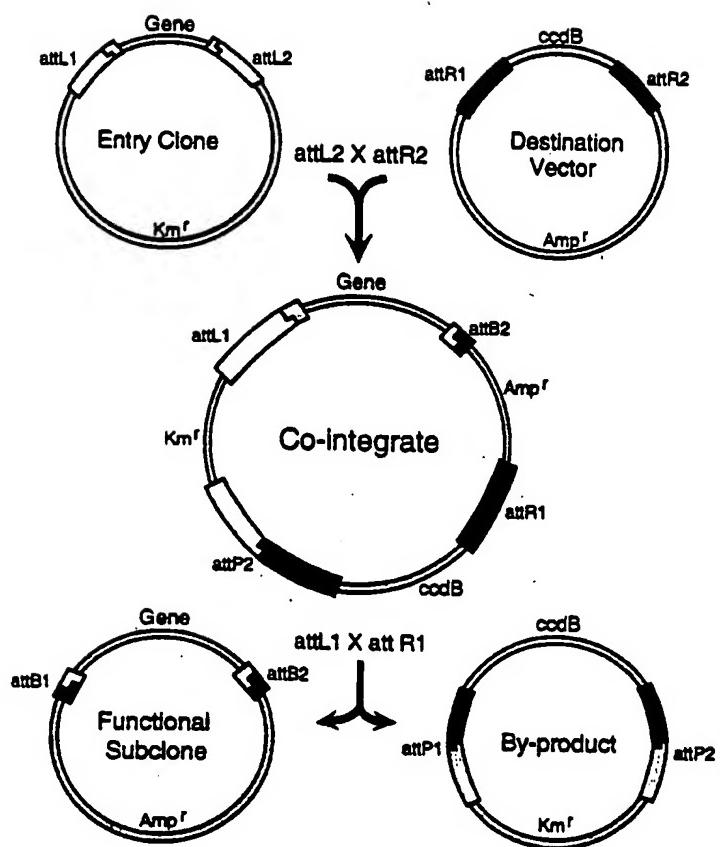


FIGURE 61

169/240

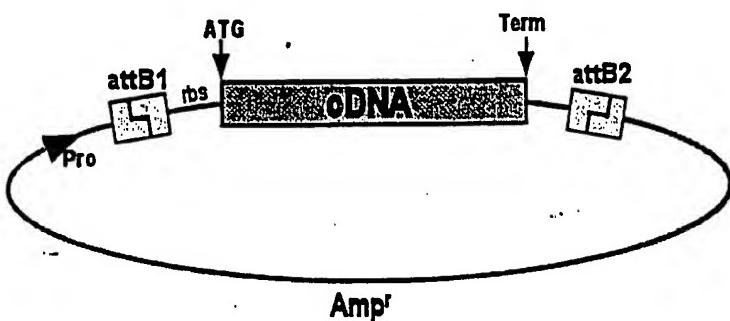
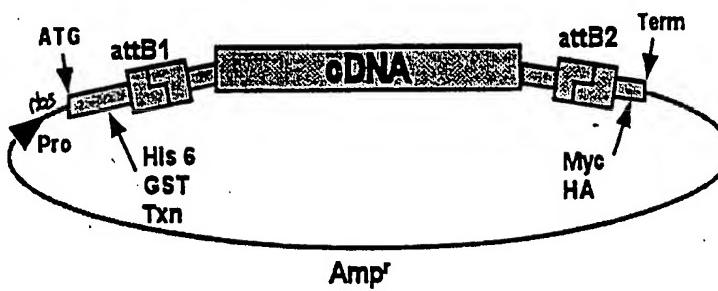
Native Protein Expression:**Fusion Protein Expression:**

FIGURE 62

170/240

Mlu I (reading frame A)

Bgl II (reading frame B)

Xba I (reading frame C)

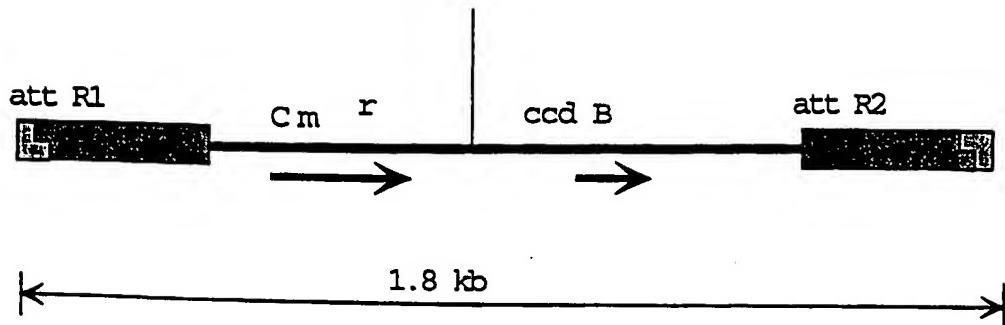


FIGURE 63

171/240

FIGURE 64A

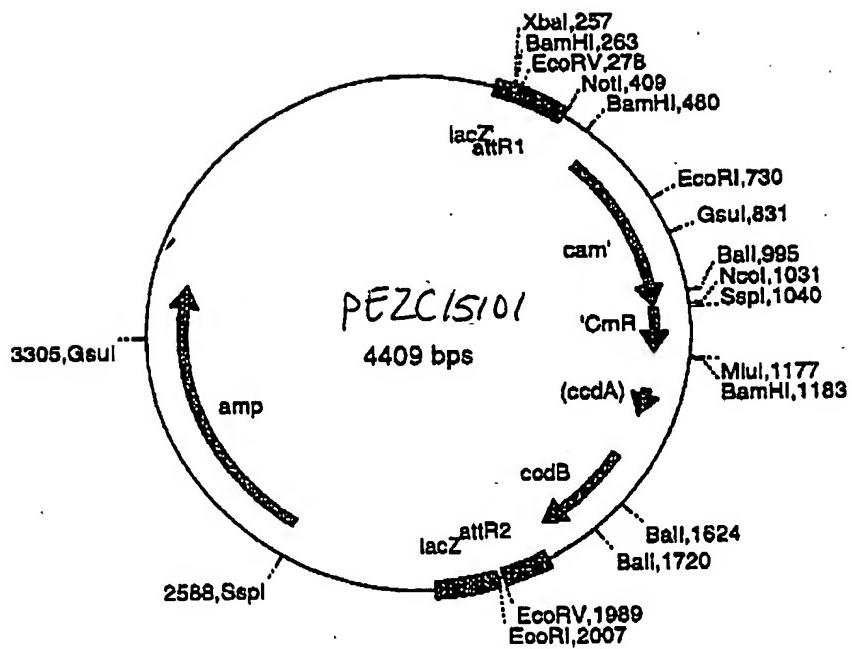
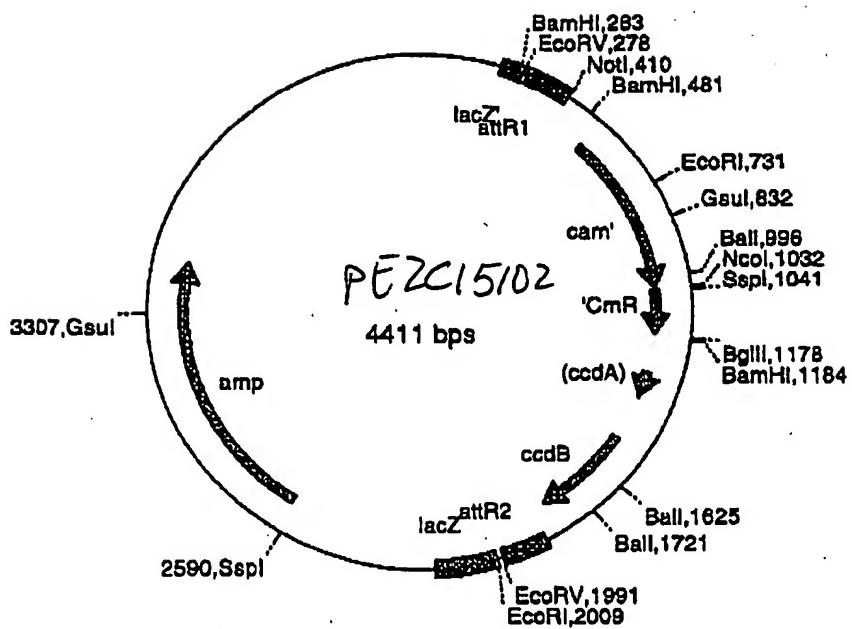


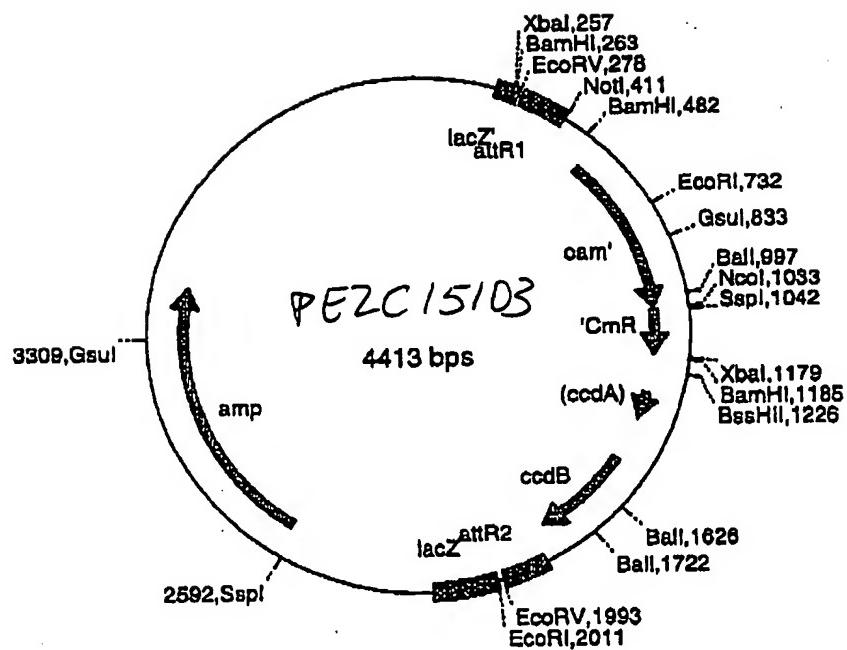
FIGURE ColB



173/240

5

FIGURE 64C



174/260

Primers for Amplifying *tetR* and *ampR*
for Cloning by Recombination

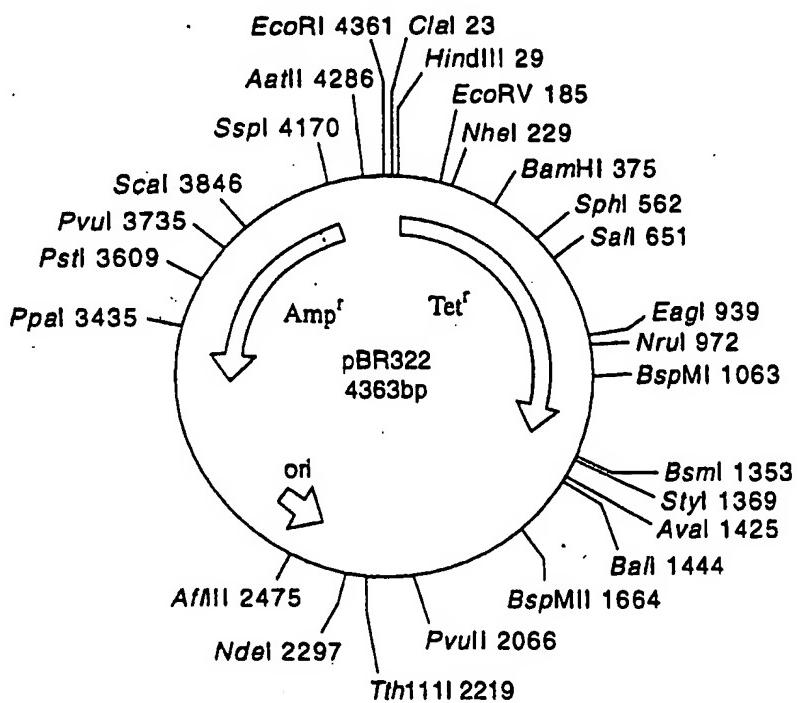
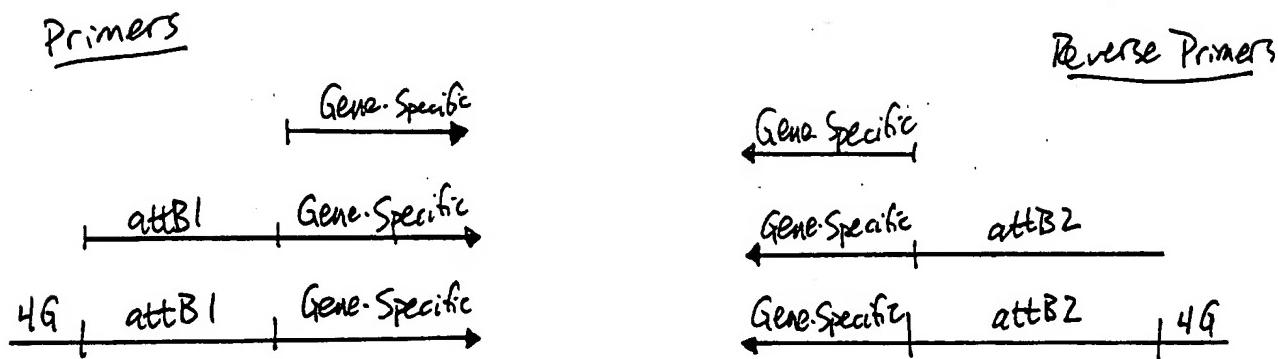


FIGURE 65

175/260

**Results of Cloning
tet and amp PCR Products
by Recombination**

PCR Product Used in GCS Reactions	No. Colonies Obtained (100 µl plated)	Form of DNA Analyzed	Colonies Obtained of Predicted Size
tet	6, 10	SC	0 of 8
attB-tet	9, 6	SC	1 of 8
attB+4G-tet	824, 1064	SC AvaI+Bam	7 of 7 7 of 7
amp	7, 13	SC	0 of 8
attB-amp	18, 22	SC	3 of 8
attB+4G-amp	3020, 3540	SC PstI	8 of 8 8 of 8
attB Plasmid (Pos. Control)	320, 394		

FIGURE 66

176/260

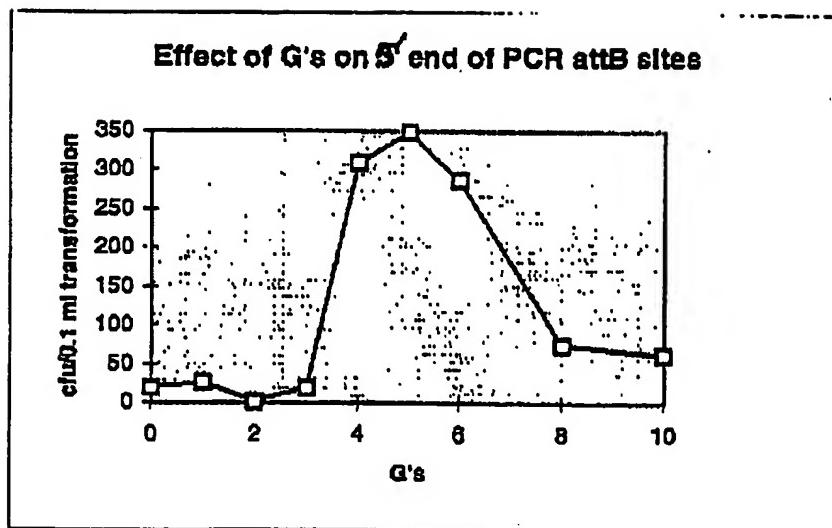


FIGURE 67

177/240

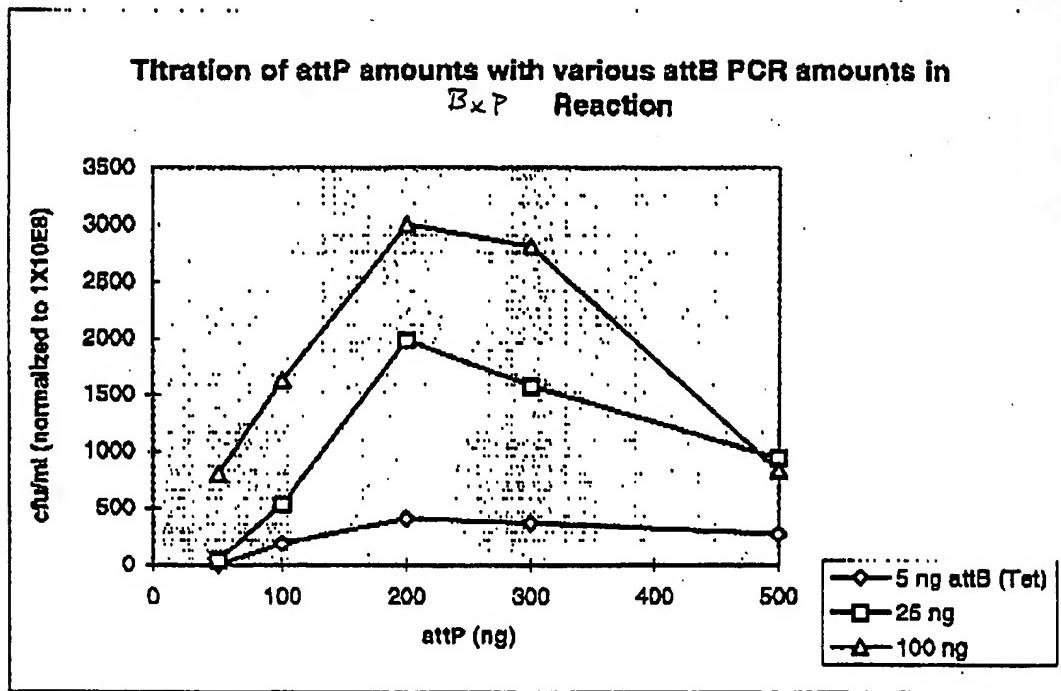
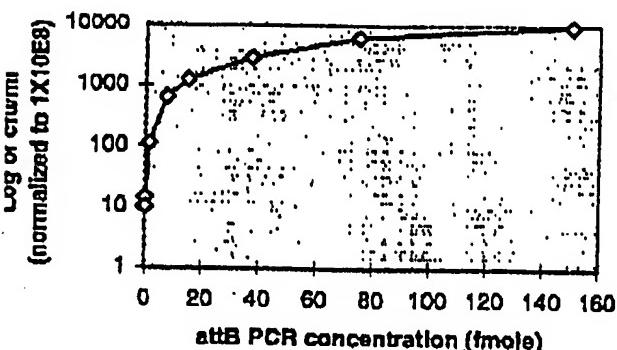
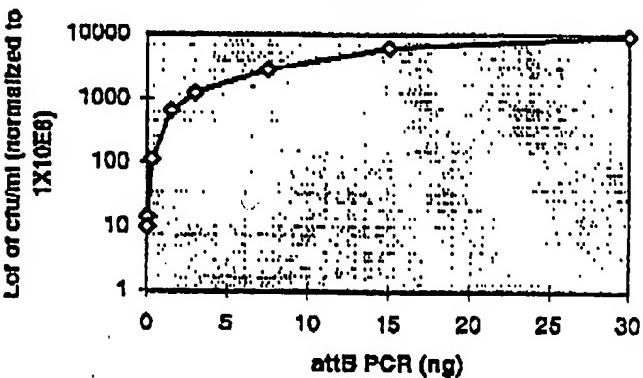
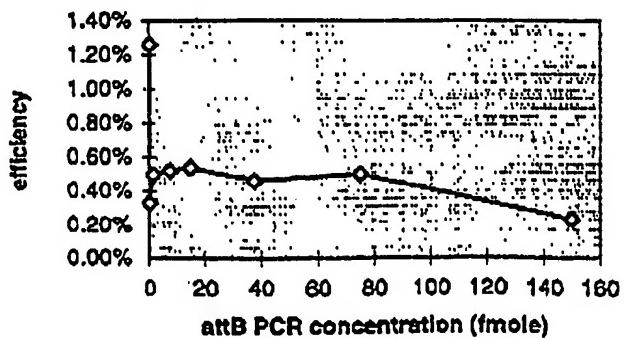


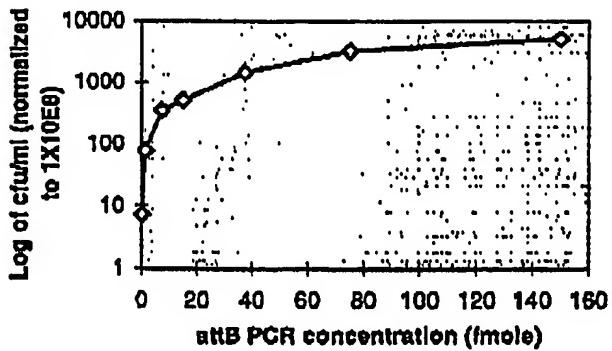
FIGURE 68

178/240

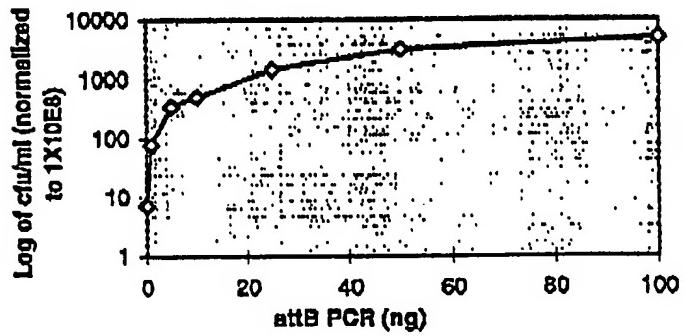
FIGURE
69**Titration of 256 bp PCR In BxP Reaction****Titration of 256 bp PCR In BxP ... Reaction****Efficiency of cloning of 256 bp PCR Into a [Destination] vector**

179/240
FIGURE
70

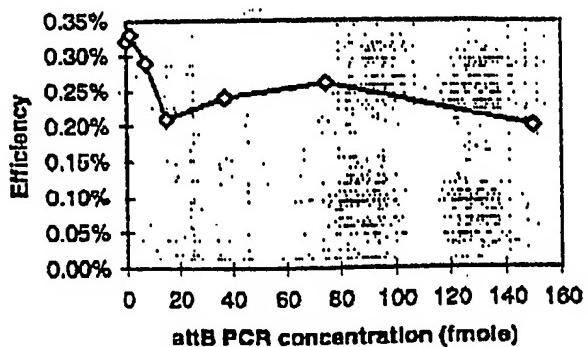
Titration of 1 kb PCR in B_{XP}
Reaction



Titration of 1 kb PCR in B_{XP}
Reaction



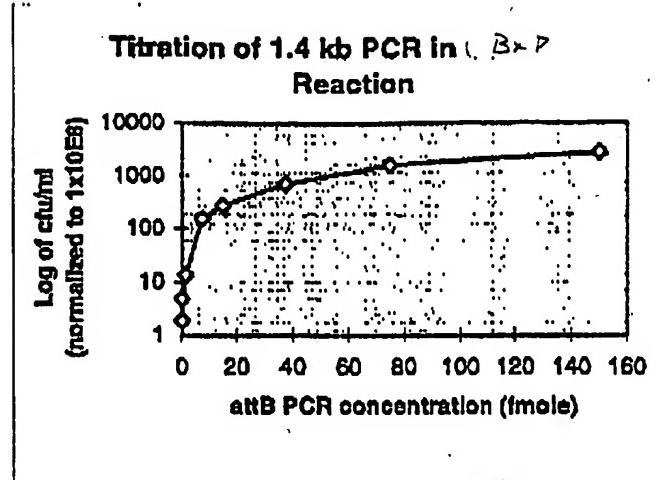
Efficiency of cloning of 1 kb PCR
into a Destination vector



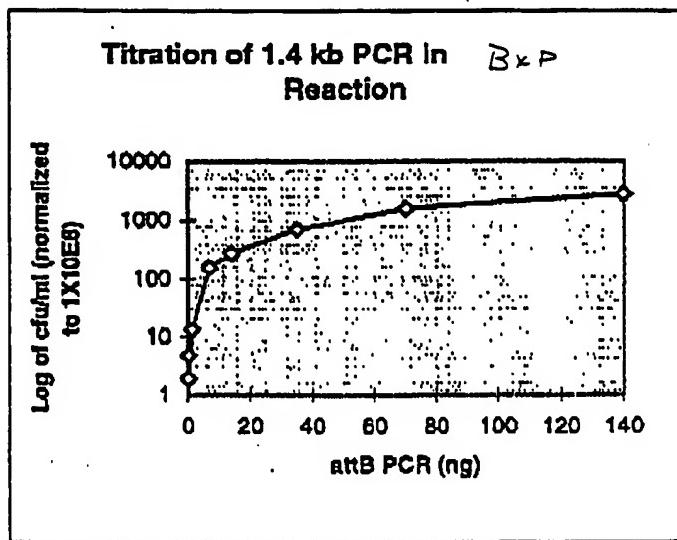
(80/260)

FIGURE 7

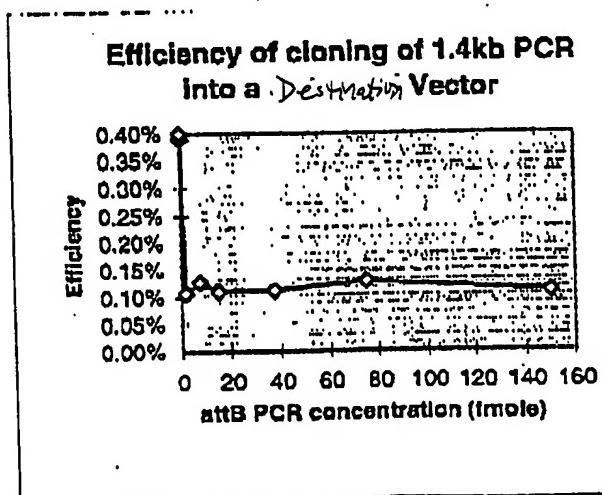
A



B



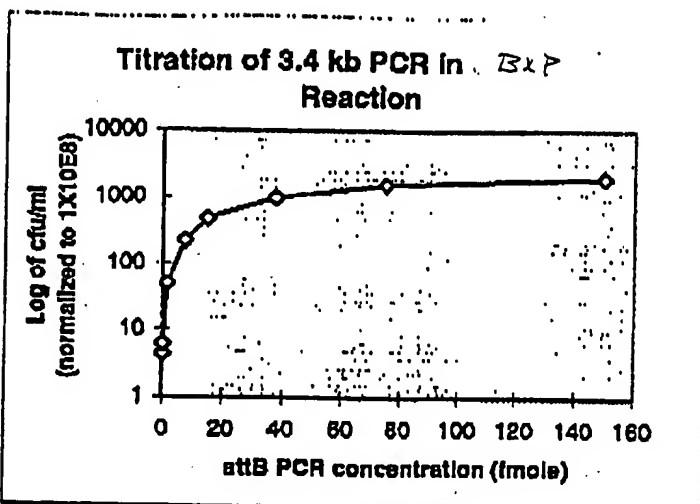
C



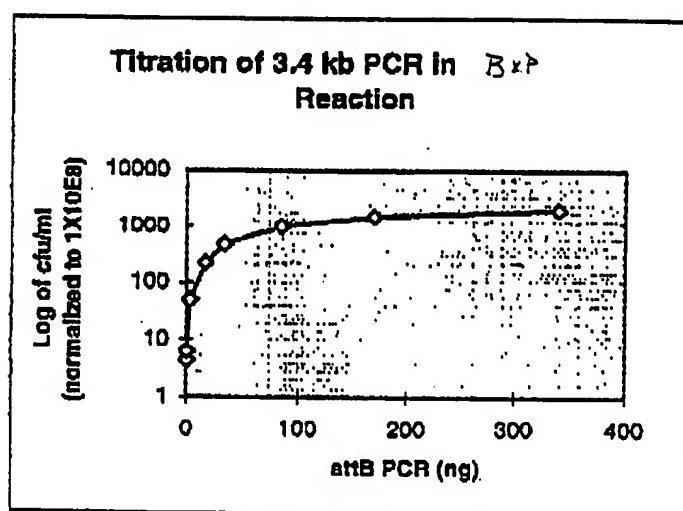
181/240

FIGURE 72

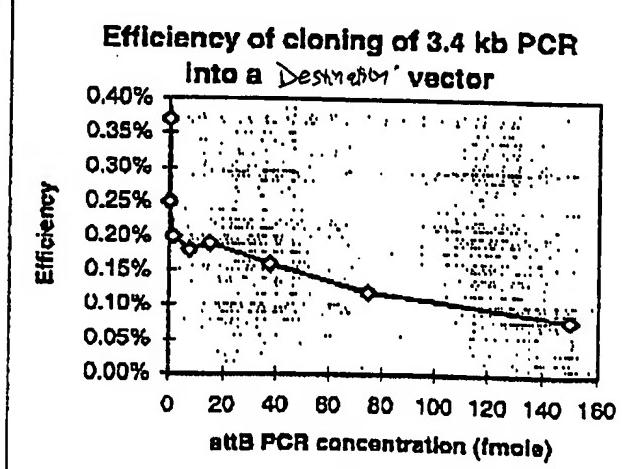
A



B



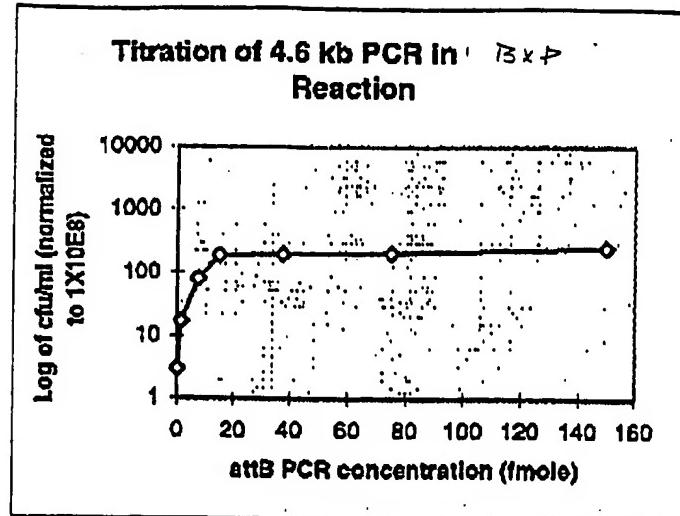
C



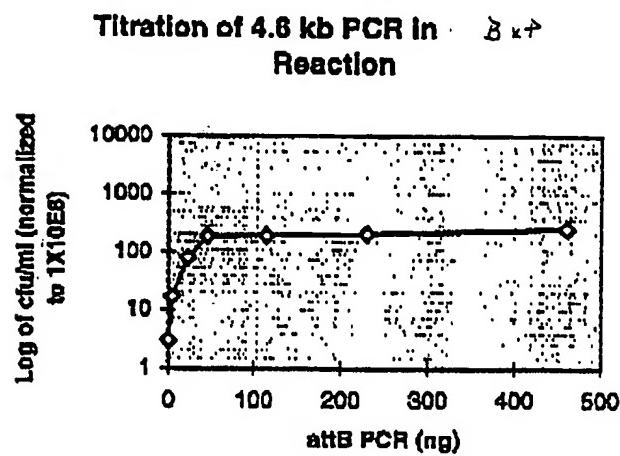
182/240

FIGURE 73

A

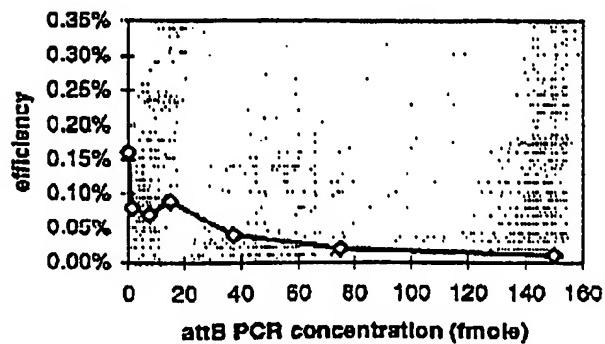


B



C

Efficiency of cloning of 4.6 kb PCR into a destination vector



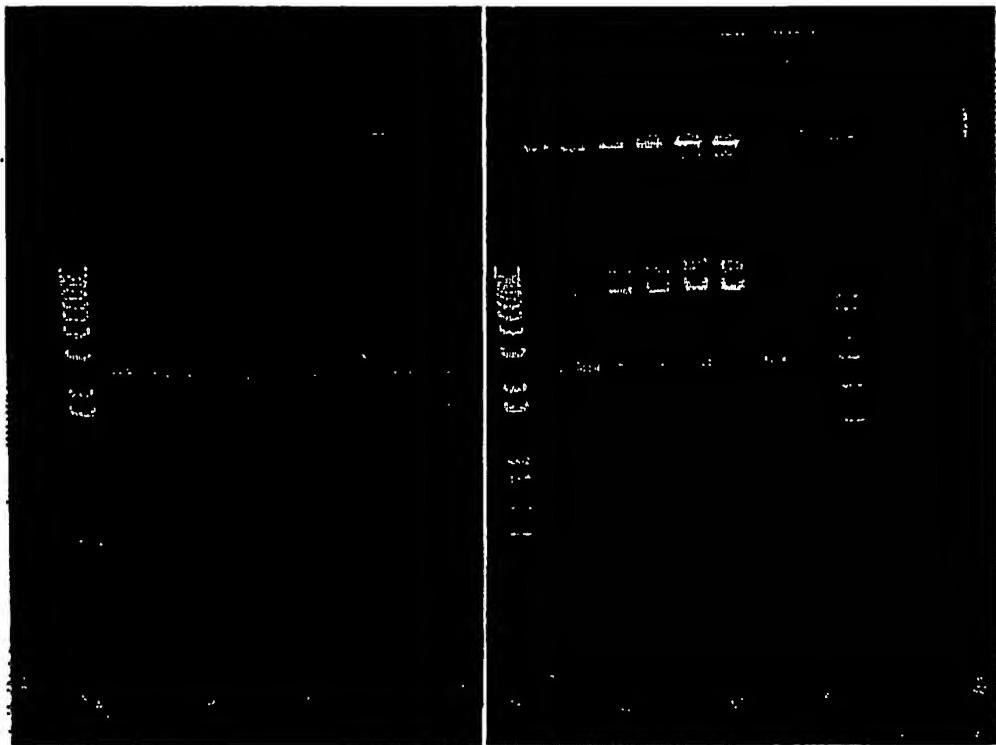
6.9 kb PCR DNA Titration in a BxP Reaction

FIGURE 74

184/240

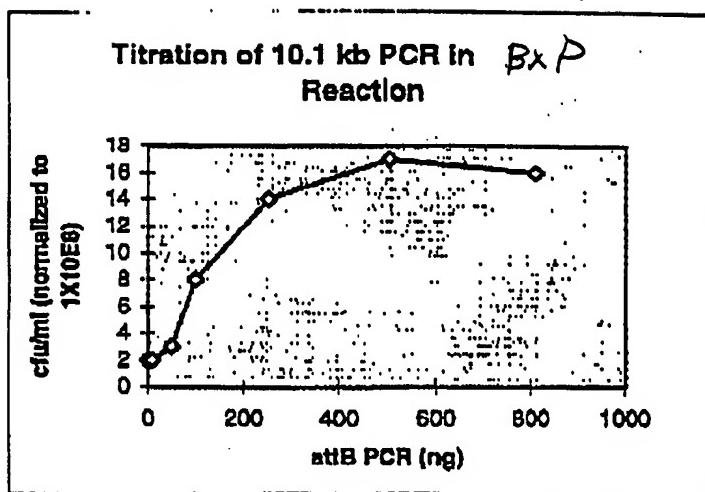


FIGURE 75-

185/240

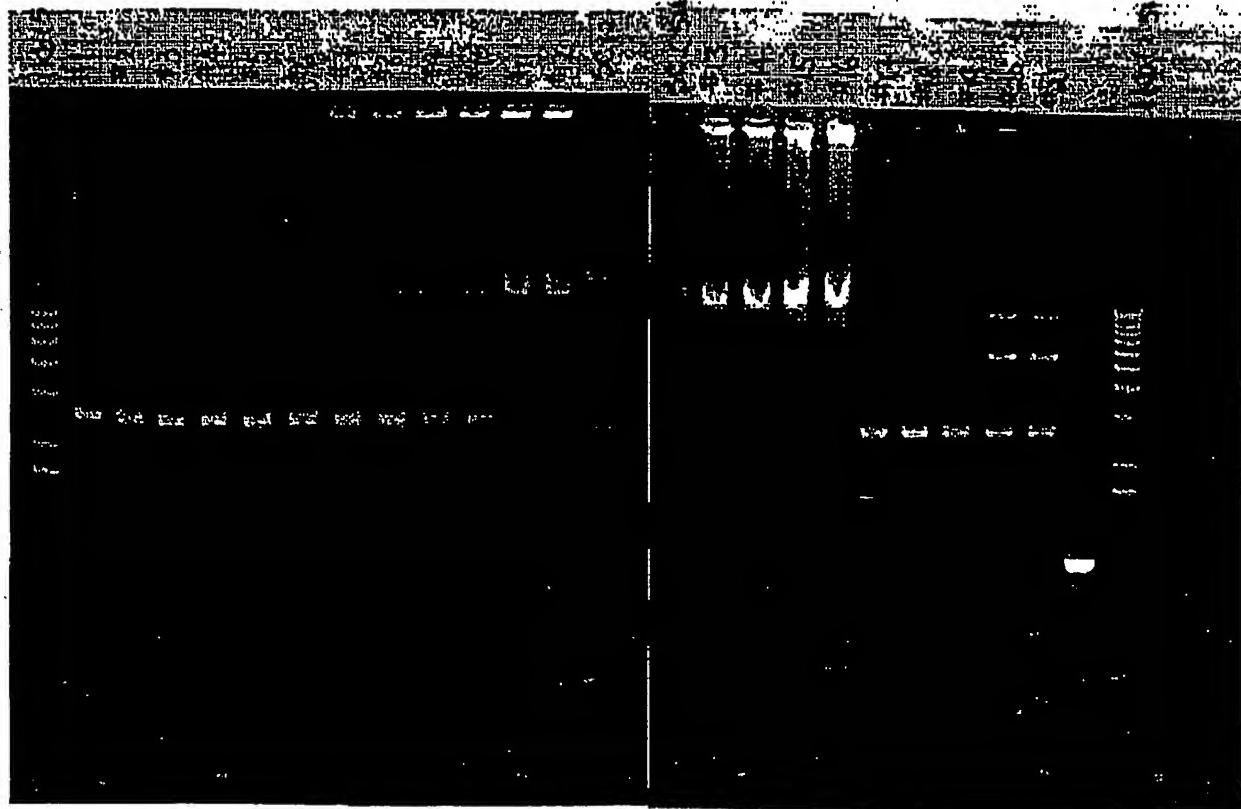
10.1 kb PCR DNA Titration in $Bx\gamma$ Reaction

FIGURE 76

186/240

**Cloning of PCR Products of Different Sizes with the
GATEWAY™ PCR Cloning System**

Size	fmols PCR DNA	ng PCR DNA	Cols/ml Transformation (pUC=10 ⁸ CFU/ml)	Correct Clones/Total Examined**
0.26 kb*	15	3	1223	10/10 (a)
	37.5	7.5	2815	
1.0 kb	15	10	507	49/50 (b)
	37.5	25	1447	
1.4 kb	15	14	271	48/50 (c)
	37.5	35	683	
3.4 kb	15	34	478	9/10 (a)
	37.5	85	976	
4.6 kb	15	46	190	10/10 (a)
	37.5	115	195	
6.9 kb	15	69	30 (235)**	47/50 (b)
	37.5	173	54 (463)**	

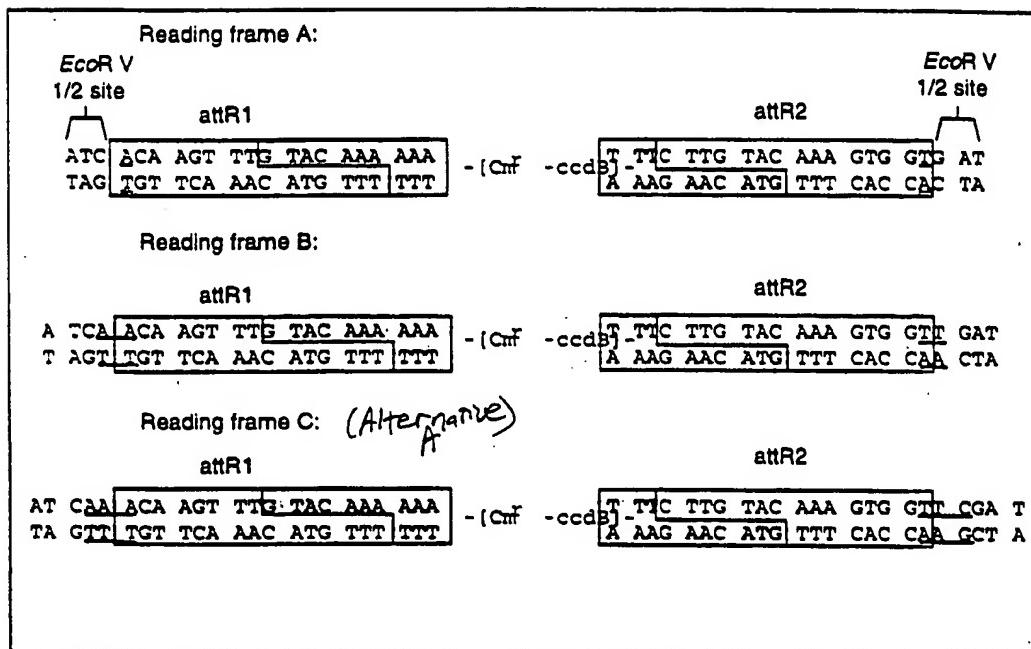
*The 0.26 kb PCR product was used unpurified; all the others were purified by precipitation with PEG/MgCl₂ as described in the text of Example 9, to remove primer dimers potentially present. Standard incubations were for 60 min.

**overnight incubation

- (a) DNA minipreps
- (b) ampR/kanR
- (c) tetR/kanR

Figure 77

187/240



Reading frame C: (Alternative)
B

FIGURE 78

188/h_o

Fusion protein

codon	Reading frame A cassette
-- nnn nnn atc aca agt ttg tac aaa aaa gct ---	attR 1
-- nnn nnn tag tgt tca aac atg ttt ttt cga ---	

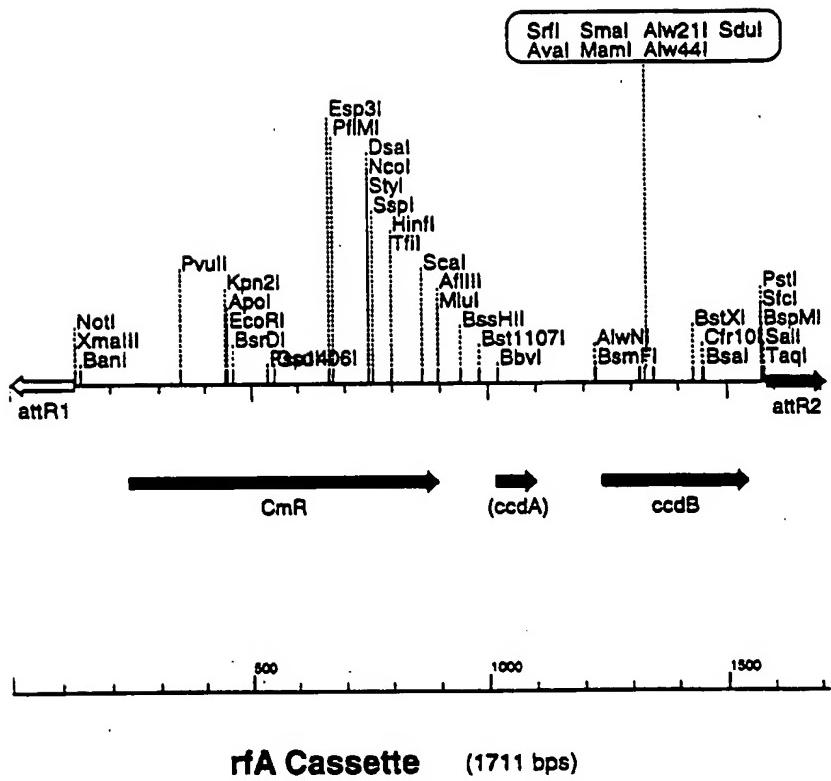
Reading frame B cassette	
-- nnn nnn nna <u>tca</u> aca agt ttg tac aaa aaa gct ---	*
-- nnn nnn nnt agt tgt tca aac atg ttt ttt cga ---	

* cannot be TG or TA

Reading frame C cassette	
-- nnn nnn nat <u>caa</u> aca agt ttg tac aaa aaa gct ---	
-- nnn nnn nta gtt tgt tca aac atg ttt ttt cga ---	

FIGURE 79

189/240



rfA Cassette (1711 bps)

190/240

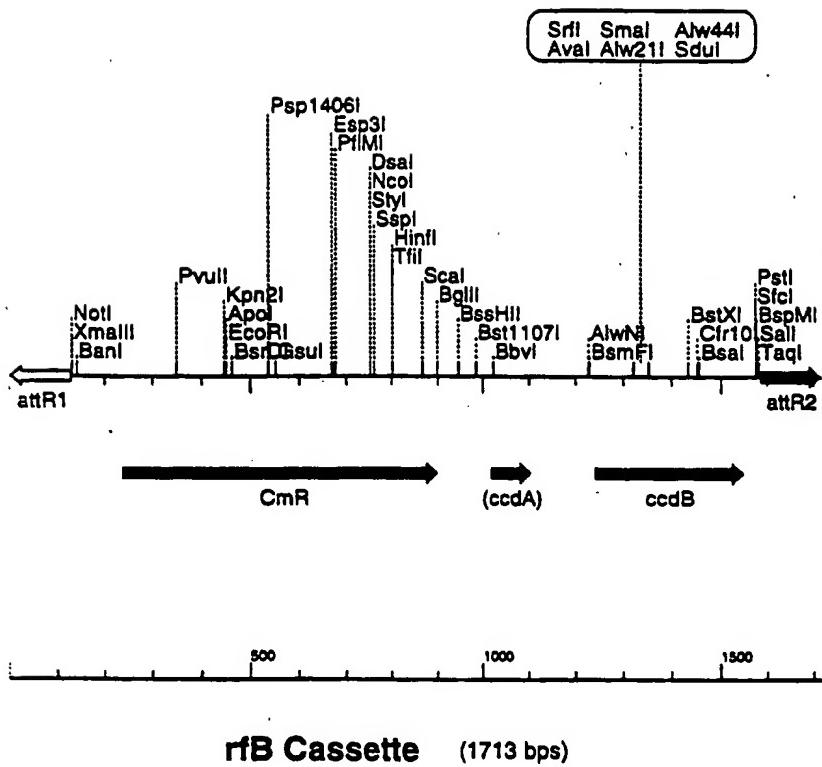
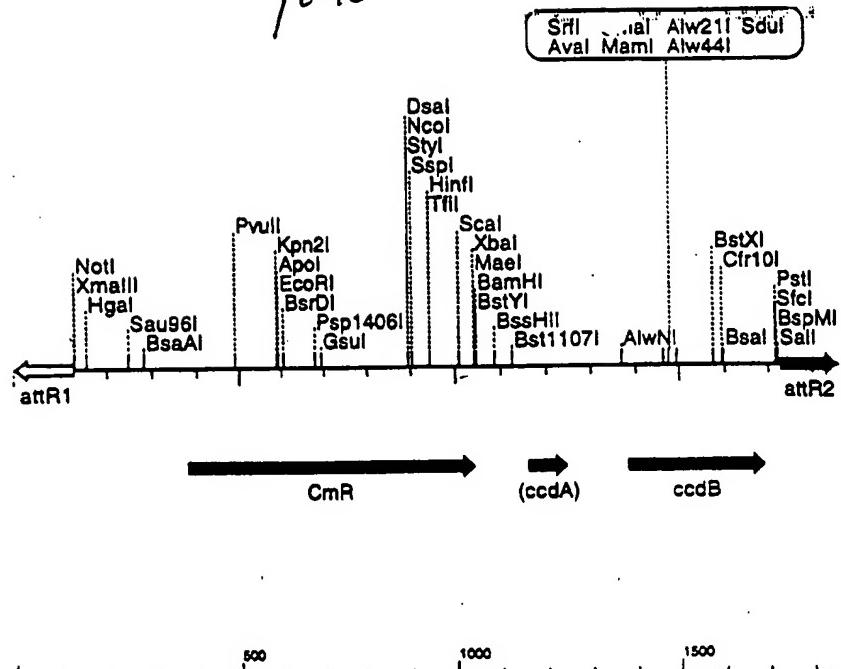
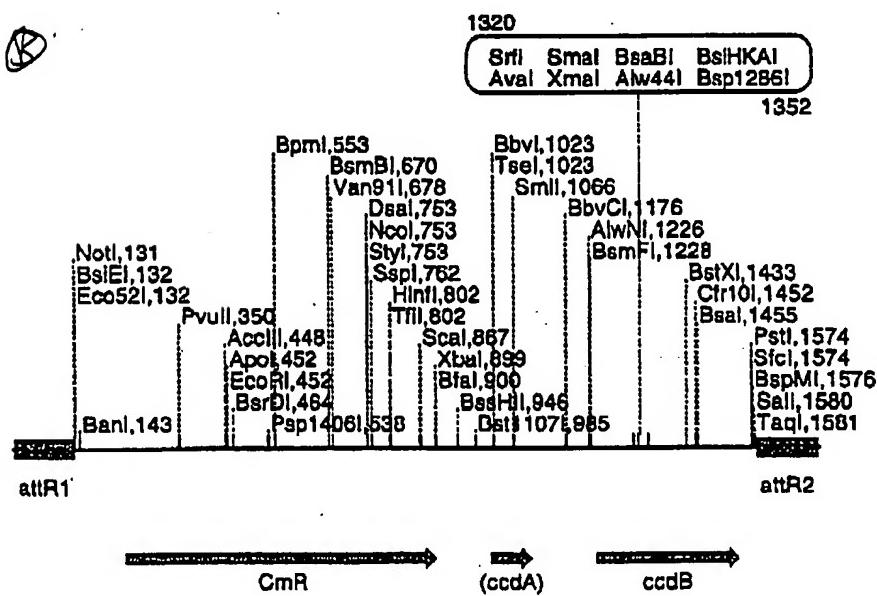


FIGURE 81

191/260



rfC Cassette



rfC cassette (1715 bps)

FIGURE 82

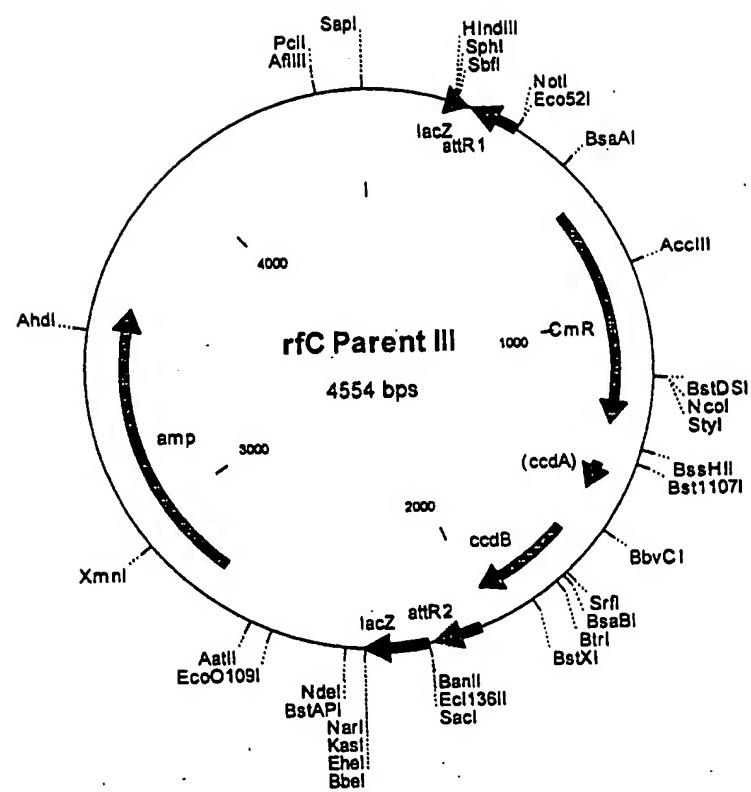


FIGURE 83 A

prfC Parent III 4554 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
410..286	attR1
660..1319	CmR
1439..1523	inactivated ccdA
1661..1966	ccdB
2007..2131	attR2
2753..3613	amp

1 GCGCCAATA CGCAAACCGC CTCTCCCCGC GCGTGGCCG ATTCATTAAT GCAGCTGGCA
 61 CGACAGTTT CCCGACTGGG AAGCGGCCAG TGAGCGAAC GCAATTAAATG TGAGTTAGCT
 121 CACTCATTAG GCACCCCAGG CTTTACACTT TATGCTCCG GCTCGTATGT TGTGTGGAAT
 181 TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC CATGATTACG CCAAGCTTGC
 241 ATGCCTGCAG GTGACTCTA GAGGATCCCC GGGTACCGAT ATCAAACAAG TTTGTACAAA
 301 AAAGCTAAC GAGAAACGTA AAATGATATA AATATCAATA TATTAATTTA GATTTTGCAT
 361 AAAAACAGA CTACATAATA CTGTAAAACA CAACATATCC AGTCACTATG GCGGCCGCTA
 421 AGTTGGCAGC ATCACCCGAC GCACCTTGCG CCGAATAAAAT ACCTGTGACG GAAGATCACT
 481 TCGCAGAATA AATAAAATCTT GGTGCTCCCTG TTGATACCGG GAAGCCCTGG GCCAACTTTT
 541 GGCAGAAATG AGACGTTGAT CGGCACGTTA GAGGTTCCAA CTTTCACCAT AATGAAATAA
 601 GATCACTACC GGGCGTATTT TTTGAGTTAT CGAGATTTTC AGGAGCTAAG GAAGCTAAAA
 661 TGGAGAAAAA AATCACTGGG TATACCAACCG TTGATATATC CCAATGGCAT CGTAAAGAAC
 721 ATTTTGAGGC ATTCAGTCA GTTGCTCAAT GTACCTATAA CCAGACCGTT CAGCTGGATA
 781 TTACGGCCTT TTTAAAGACC GTAAAGAAAA ATAAGCACAA GTTTTATCCG GCCTTTATT
 841 ACATTCTTGC CGGCCTGATG AATGCTCATC CGGAATTCCG TATGGCAATG AAAGACGGTG
 901 AGCTGGTGAT ATGGGATAGT GTTCACCCCTT GTTACACCGT TTTCCATGAG CAAACTGAAA
 961 CGTTTCATC GCTCTGGAGT GAATACACG ACGATTTCCG GCAGTTCTA CACATATATT
 1021 CGCAAGATGT GGCGTGTAC GGTGAAACCC TGGCCTATTT CCCTAAAGGG TTTATTGAGA
 1081 ATATGTTTTT CGTCTCAGCC AATCCCTGGG TGAGTTTCAC CAGTTTTGAT TTAAACGTGG
 1141 CCAATATGGG CAACTTCTTC GCCCCCGTT TCACCATGGG CAAATATTAT ACGCAAGGCG
 1201 ACAAGGTGCT GATGCCGCTG GCGATTCCAGG TTCATCATGC CGTCTGTGAT GGCTTCCATG
 1261 TCGGCAGAAT GTTTAATGAA TTACAAACAGT ACTGCGATGA GTGGCAGGGC GGGGCGTAAT
 1321 CTAGAGGATC CGGCTTACTA AAAGCCAGAT AACAGTATGC GTATTTGCGC GCTGATTTTT
 1381 GCGGTATAAG AATATATACT GATATGTATA CCCGAAGTAT GTCAAAAGA GGTGTGCTAT
 1441 GAAGCAGCGT ATTACAGTGA CAGTTGACAG CGACAGCTAT CAGTTGCTA AGGCATATAT
 1501 GATGCAATA TCTCCGGTCT GGTAAAGACA ACCATGCAGA ATGAAGCCCG TCGCTGCGT
 1561 GCCGAACGCT GGAAAGCGGA AAATCAGGAA GGGATGGCTG AGGTCGCCCG GTTTATTGAA
 1621 ATGAACGGCT CTTTGCTGA CGAGAACAGG GACTGGTGA ATGCAGTTA AGGTTACAC
 1681 CTATAAAAGA GAGAGCCGTT ATCGTCTGTT TGTTGATGTA CAGAGTGATA TTATTGACAC
 1741 GCCCCGGCGA CGGATGGTGA TCCCCCTGGC CAGTGCACGT CTGCTGTCAAG ATAAAGTCTC
 1801 CCGTGAACCT TACCCGGTGG TGCATATCGG GGATGAAAGC TGGCGCATGA TGACCACCGA
 1861 TATGCCAGT GTGCCGGTCT CCGTTATCGG GGAAGAAGTG GCTGATCTA GCCACCGCGA
 1921 AAATGACATC AAAACGCCA TTAACTGAT GTTCTGGGA ATATAAAATG CAGGCTCCGT
 1981 TATACACAGC CAGTCTGCAG GTGACCCATA GTGACTGGAT ATGTTGTGTT TTACAGTATT
 2041 ATGTAGTCTG TTTTTTATGC AAAATCTAAT TTAATATATT GATATTTATA TCATTTCAGC
 2101 TTTCTCGTTC AGCTTTCTTG TACAAAGTGG TTGATATCG GTACCGAGCT CGAATTCACT
 2161 GGCGTCGTT TTACAACGTC GTGACTGGGA AAACCCCTGGC GTTACCCAC TTAATCGCCT
 2221 TGCAGCACAT CCCCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC
 2281 TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCCTG ATGCGGTATT TTCTCCTTAC
 2341 GCATCTGTGC GGTATTTCAC ACCGCATATG GTGCACTCTC AGTACAATCT GCTCTGATGC
 2401 CGCATAGTTA AGCCAGCCCC GACACCCGCC AACACCCGCT GACGCCCT GACGGGCTTG
 2461 TCTGCTCCCG GCATCCGCTT ACAGACAAAGC TGTGACCGTC TCCGGGAGCT GCATGTGTCA
 2521 GAGGTTTTCA CGTCATCAC CGAAACCGCG GAGACGAAAG GGCCTCGTGA TACGCCATT
 2581 TTATAGGTT AATGTCATGA TAATAATGGT TTCTTAGACG TCAGGTGGCA CTTTCCGGG
 2641 AAATGTCGCG GGAACCCCTA TTGTTTATT TTCTAAATA CATTCAAATA TGATCCGCT
 2701 CATGAGACAA TAACCCCTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT
 2761 TCAACATTTC CGTGTGCCCTT TTATCCCTT TTTGCGGCCA TTTTGCCCTC CTGTTTTGCA-

FIGURE 83B

2821 TCACCCAGAA ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG
2881 TTACATCGAA CTGGATCTCA ACAGCGTAA GATCCTTGAG AGTTTTCGCC CGGAAGAACG
2941 TTTTCCAATG ATGAGCACTT TTAAAGTCT GCTATGTGGC GCGGTATTAT CCCGTATTGA
3001 CGCCCCGCAA GAGCAACTCG GTCCGCGCAT ACACATTCT CAGAATGACT TGGTTGAGTA
3061 CTCACCAGTC ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAT TATGCAGTGC
3121 TGCCATAACC ATGAGTGATA ACACGTGGC CAACTTACTT CTGACAAACGA TCGGAGGACC
3181 GAAGGAGCTA ACCGCTTTTT TGCAACACAT GGGGGATCAT GTAACTCGCC TTGATCGTTG
3241 GGAACCGGAG CTGAATGAAG CCATACAAA CGACGAGCGT GACACCACGA TGCCTGTAGC
3301 AATGGCAACA ACGTTGCGCA AACTATTAAAC TGGCGAACTA CTTACTCTAG CTTCCCGCA
3361 ACAATTAAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA CCACTTCTGC GCTCGCCCT
3421 TCCGGCTGGC TGGTTTATTG CTGATAAAATC TGGAGCCGGT GAGCGTGGGT CTCGCGGTAT
3481 CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCCTATC GTAGTTATCT ACACGACGGG
3541 GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
3601 TAAGCATTGG TAACTGTCAG ACCAAGTTA CTCATATATA CTTTAGATTG ATTTAAAATC
3661 TCATTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACAAAAT
3721 CCCTTAACGT GAGTTTTCGT TCCACTGAGC GTCAAGACCCC GTAGAAAAGA TCAAAGGATC
3781 TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCAACGCT
3841 ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACGG
3901 CTTCAGCAGA GCGCAGATAAC CAAATACTGT CCTTCTAGTG TAGCCGTAGT TAGGCCACCA
3961 CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC
4021 TGCTGCCAGT GGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA
4081 TAAGGCGCAG CGGTGCGGGCT GAACGGGGGG TTGGTGCACA CAGCCCAGCT TGGAGCGAAC
4141 GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCGA
4201 AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG
4261 GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCT GTGGGGTTTC GCCACCTCTG
4321 ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGGCGG AGCCTATGGA AAAACGCCAG
4381 CAACGCGGCC TTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTCC
4441 TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC
4501 TCGCCGCAGC CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGA

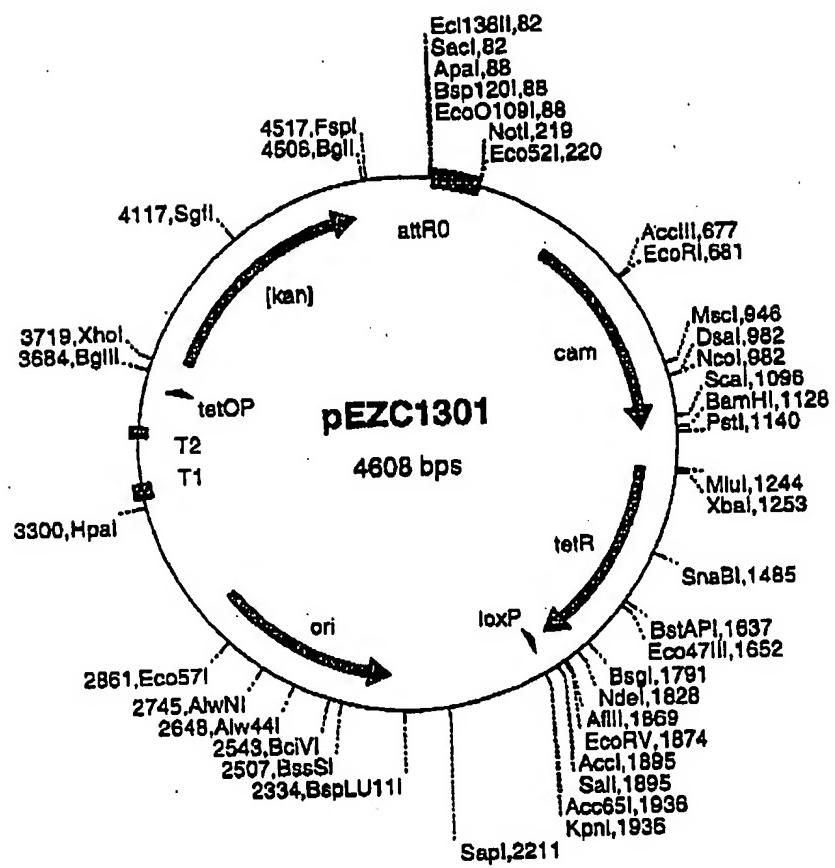


FIGURE 84

196/260

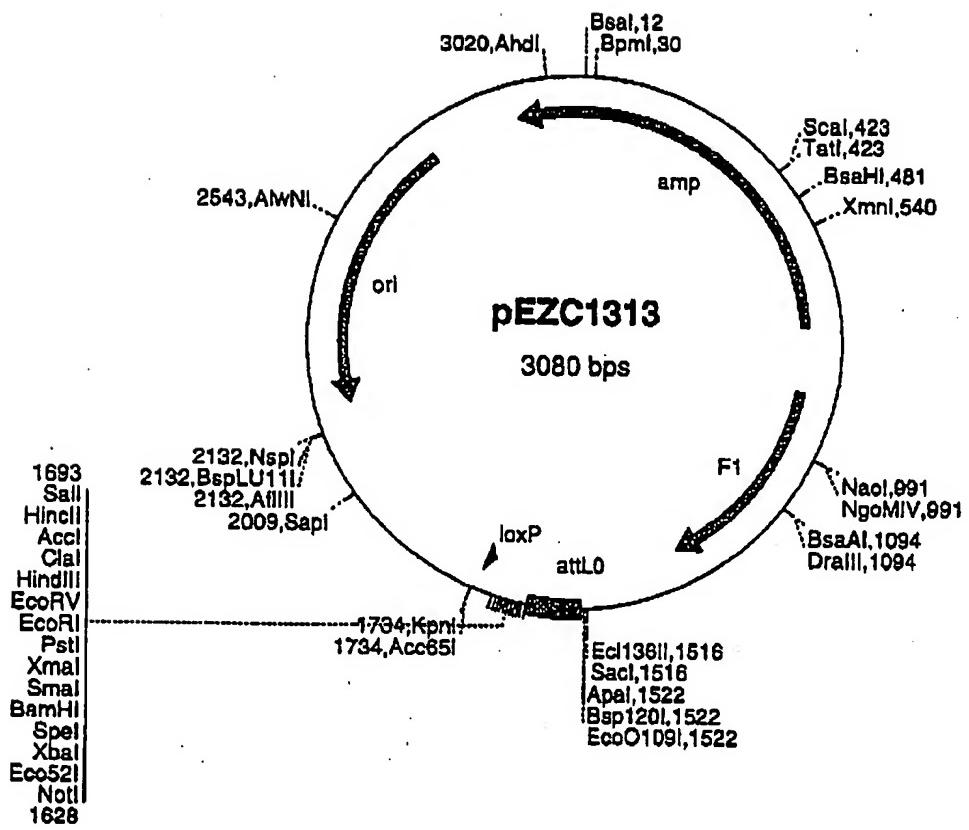


FIGURE 85

97/240

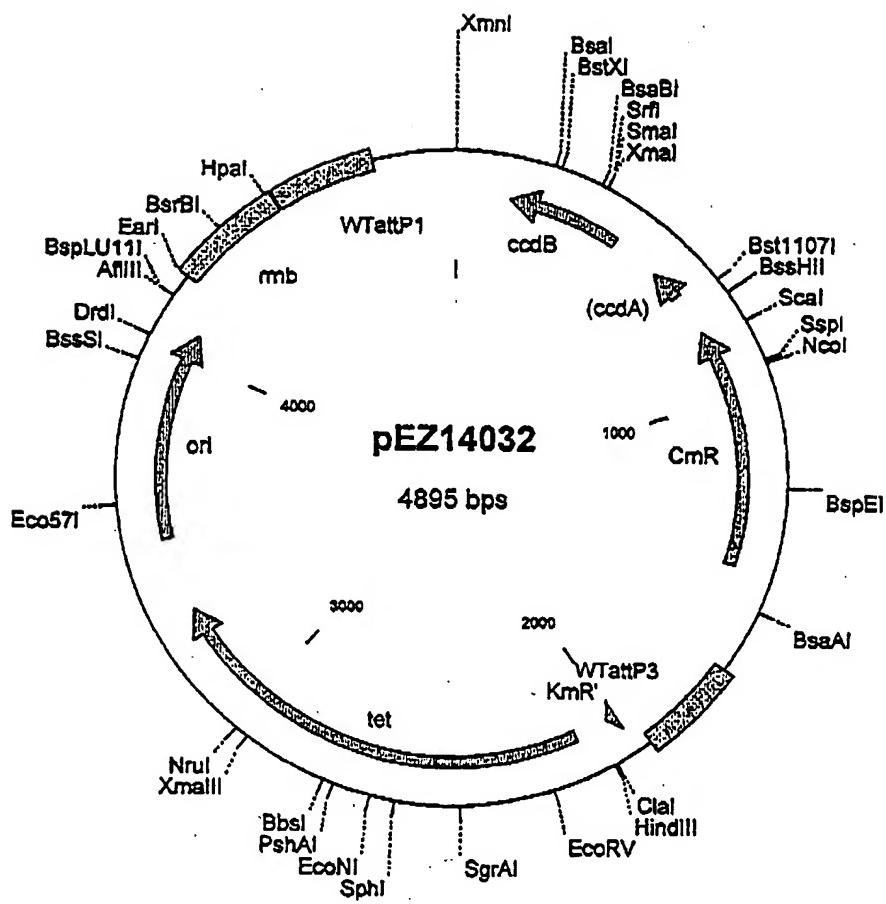


FIGURE 86

198/260

FIGURE 87

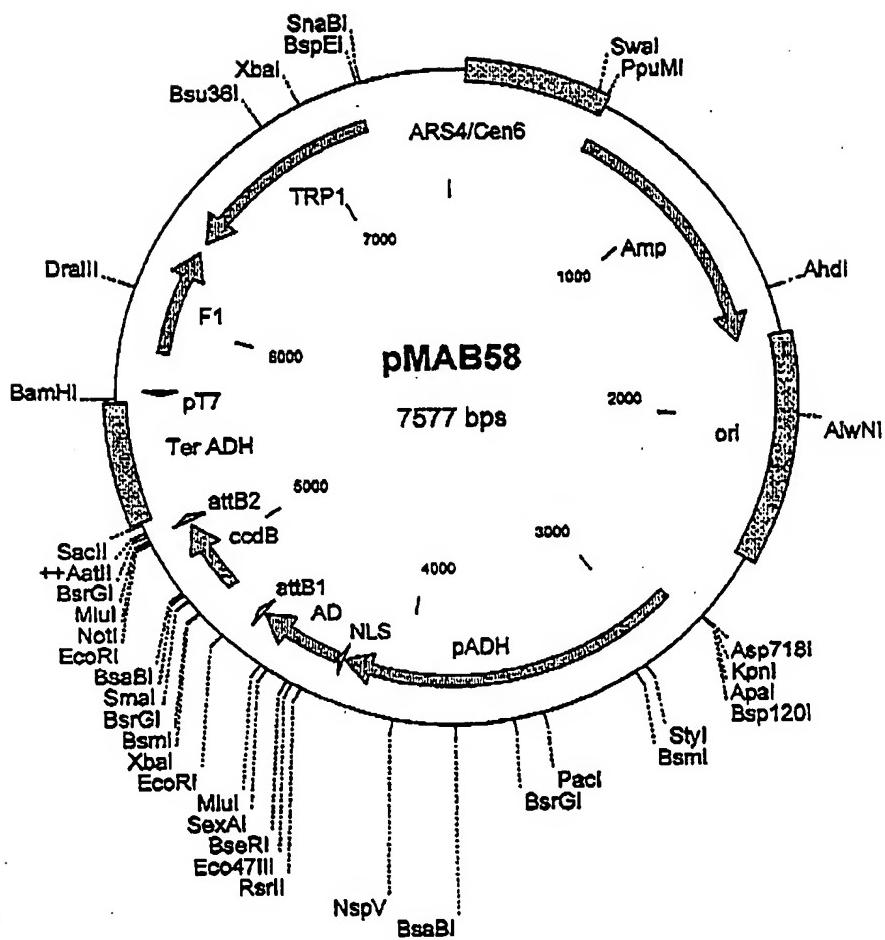
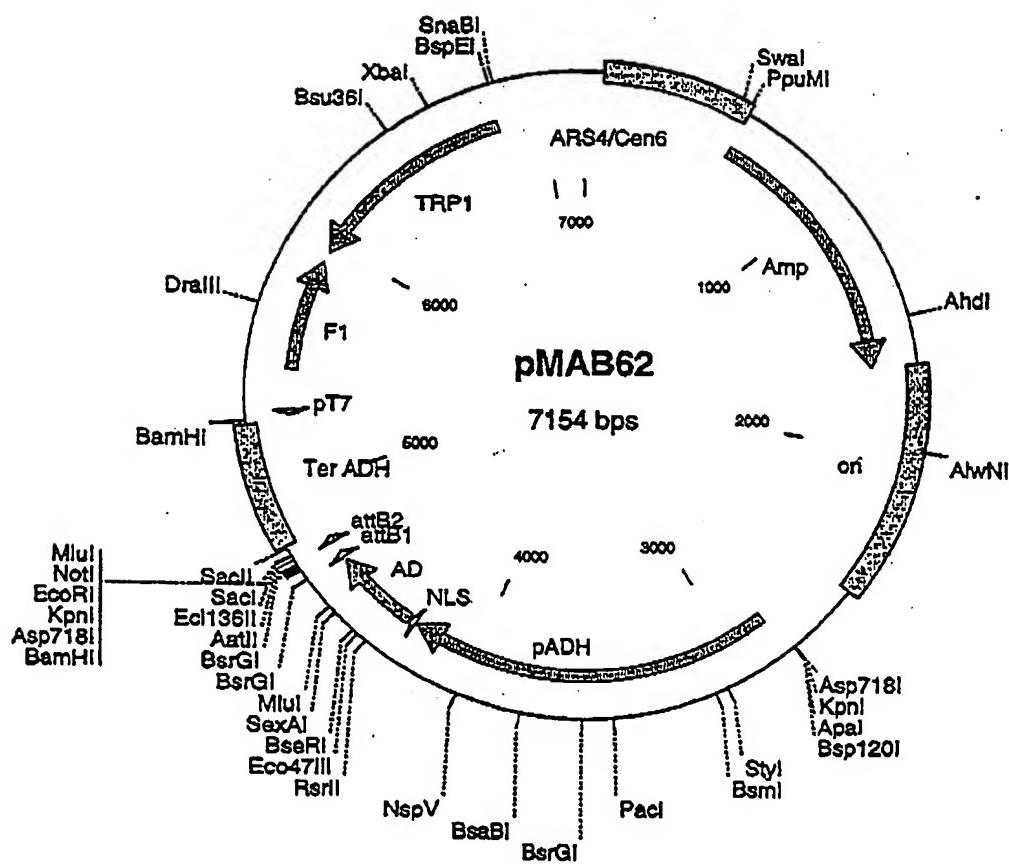


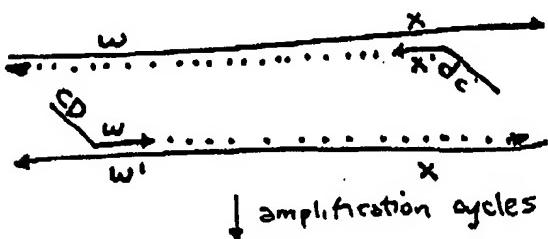
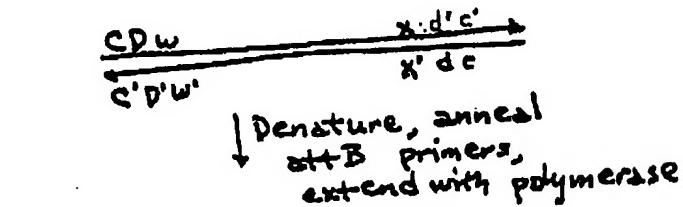
FIGURE 88



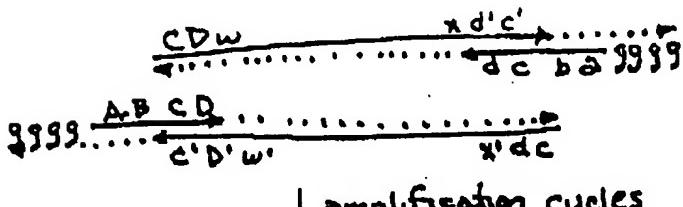
200/240

DNA to be amplified ($5' \rightarrow 3'$):

\downarrow Denature, anneal
hybrid primers,
 \downarrow extend with polymerase

 \downarrow amplification cycles

\downarrow Denature, anneal
 \downarrow attB primers,
extend with polymerase

 \downarrow amplification cycles

attB1 primer:
gggg \xrightarrow{ABCD}

attB2 primer:
gggg \xrightarrow{abcd}

Hybrid primers (part attB, part gene specific):

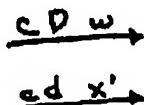


FIGURE 89

201/240

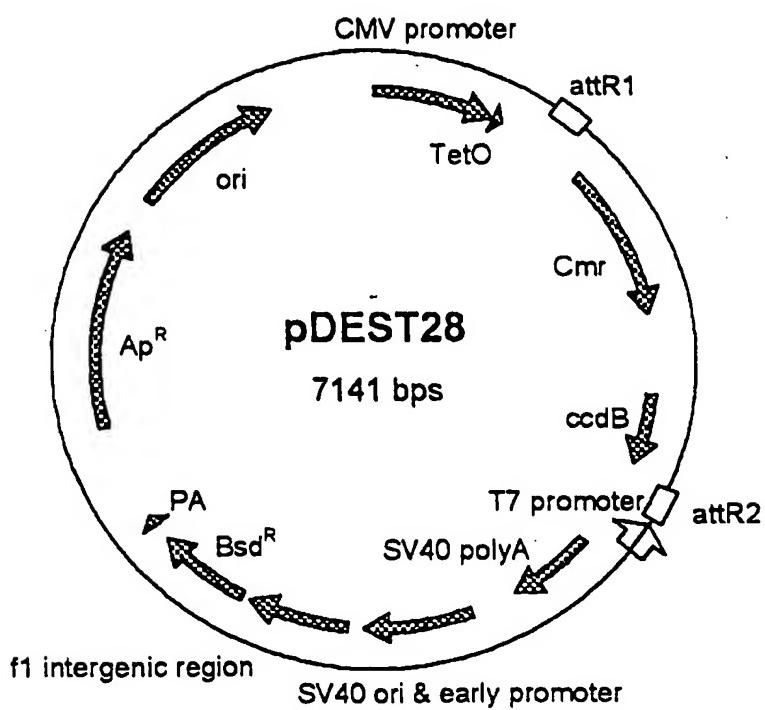


FIGURE 90A

202/240

PDEST28 7141 bp

ATGCATGTCGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCC
 CGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCAT
 TGACGTCAATGGGTGGAGTATTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTAT
 CATATGCCAAGTACGCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTAT
 GCCCAGTACATGACCTTATGGGACTTCTACTTGGCAGTACATCTACGTATTAGTCATC
 GCTATTACCATGGTGATGCGGTTGGCAGTACATCAATGGCGTGGATAGCGGTTGAC
 TCACGGGATTCCAAGTCTCACCCCATGACGTCAATGGGAGTTGTTTGGCACCAA
 AATCAACGGGACTTCCAAAATGTCGTAACAACCTCCGCCCCATTGACGCAAATGGCGGT
 AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCAGTGTAGAGATCTC
 CCTATCAGTGTAGAGATCGTCGACGAGCTGTTAGTGAACCGTCAGATCGCTGGAGA
 CGCCATCCACGCTTTGACCTCCATAGAACACCCGGACCGATCCAGCCTCCGACT
 CTAGAGGATCCCTACCGGTGATATCCTCGAGCCCATCAACAAGTTGTACAAAAAGCTG
 AACGAGAAACGTAAAATGATATAAAATCAATATAAAATTAGATTTGCATAAAAAC
 AGACTACATAACTGTAAAACACAATATCCAGTCAGTGTGGATTTGAGTTAGGATCC
 CCCAGGTTTACACTTATGCTCCGGCTCGTAAATGTGTGGATTTGAGTTAGGATCC
 GGCGAGATTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGATATACCAC
 CGTTGATATATCCCAATGGCATCGTAAAGAACATTTCAGTCAGTGTGGCTCA
 ATGTACCTATAACCAAGACCGTTAGCTGGATATTACGGCCTTTAAAGACCGTAAAGAA
 AAATAAGCACAAGTTTATCCGGCTTATTACACATTCTGCCCGCTGATGAATGCTCA
 TCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGTATGGGATAGTGTACCC
 TTGTTACACCGTTTCCATGAGCAAACGTTTCTACGCTCTGGAGTGAATACCA
 CGACGATTCCGGCAGTTCTACACATATTCGCAAGATGTGGCTGTACGGTAAAAA
 CCTGGCCTATTCCCTAAAGGGTTATTGAGAATATGTTTCTGCTCAGCCAATCCCTG
 GGTGAGTTTACACAGTTTGTGTTAACGTTGCAATATGGACAACCTCTTCGCCCCGT
 TTTCACCATGGGAAATATTACGCAAGGCAGACAGGTGCTGATGCCGCTGGCGATTCA
 GGTTCATCATGCCGCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACA
 GTACTGCGATGAGTGGCAGGGCGGGCGTAAAGATCTGGATCCGGTTACTAAAGCCAG
 ATAACAGTATGCGTATTCGCGCTGATTTGCGGTATAAGAATATACTGATATGTA
 TACCCGAAGTATGTCAAAAGAGGTGCTATGAAGCAGCGTATTACAGTGTACAGTTGAC
 AGCGACAGCTATCAGTTGCTCAAGGCATATGATGTCATATCTCCGGTCTGGTAAGCA
 CAACCATGCGAATGAAGCCGTCGTCTCGTGCACGCTGGAAAGCGGAAATCAGG
 AAGGGATGGCTGAGGTGCCCCGTTATTGAAATGAACGGCTTTGCTGACCGAGAAC
 GGGACTGGTGAATGCAAGTTAACGTTACACCTATAAAAGAGAGAGGCCGTTATCGTCTG
 TTTGTGGATGTACAGAGTGTGATATTGACACGCCGGCGACGGATGGTGTACCCCTG
 GCCAGTGCACGTCTGCTGTCAGATAAGCTCCCGTGAACCTTACCCGGTGGTCATATC
 GGGGATGAAAGCTGGCGATGACGACCCGATATGCCAGTGTGCCGGTCTCCGGTATC
 GGGGAAGAAGTGGCTGATCTAGCCACCGCGAAAATGACATCAAAACGCCATTAAACCTG
 ATGTTCTGGGAATATAATGTCAGGCTCCCTATACACAGCCAGTGTGAGTCGACCA
 TAGTGACTGGATATGTTGTTACAGTATTATGTTGCTGTTTATGCAAATCTA
 ATTAAATATATTGATATTATCATTACGTTCTCGTCAGCTTCTGTAACAAAGT
 GGTTGATGGCGCCGCTTAGAGGGCCCAAGCTTACGCGTGATGCGACGTCAAGCTC
 TCTCCCTATAGTGAGTCGTTATAAGCTAGGCACTGCCGCTGTTTACACGTCGAT
 CTGGGAAACTGCTAGCTTGGATCTTGTGAGGAACCTTACTCTGTTGTCGACATA
 ATTGGACAAACTACAGAGATTTAAAGCTTAAGGTAATATAAAATTGTTAAGTGT
 ATAATGTTGTTAAACTAGCTGCATATGCTTGTGCTGAGAGTTGCTTACTGAGTATGA
 TTATGAAAATATTACACAGGAGCTAGTGTGTTCTGTTGAGGAAACCTTACTCTGTTG
 CAGTCCCAAGGCTCATTTCAGGCCCTCAGCTCACAGTGTGTTCTGATGATCATAATCAG
 CCATACACATTGAGAGGTTTACTGCTTAAAAAACCTCCACACCTCCCGTGA
 CCTGAAACATAAAATGAATGCAATTGTTGTTAACTGTTTATTGCAAGCTTATAATGG
 TTACAAATAAAAGCAATAGCATCACAAATTTCAAAATAAGCATTGTTACTGCTTC
 TAGTTGTGGTTTGTCCAAACTCATCAATGTTCTTATCATGTCGGATCGATCCTGCA
 AATGAATCGGCCAACCGCGGGGAGAGGCCGTTTGCAGTGTGGCTGGCTGAAAGCG
 AGGCCCGCACCAGTCGCCCTTCCAAACAGTTGCGCAGCCTGAAATGGCAATGGGAC
 CCTGTAGCGGGCATTAAAGCGCGGGTGTGGTTACGCGCAGCGTACCGCTACAC
 TTGCCAGCGCCCTAGCGCCGCTCCTTCGCTTCTCCCTTCTGCCACGTTCG
 CGGGCTTCCCCGTCAGCTAAATCGGGGCTCCCTTAGGGTCCGATTAGTGTCTT-

FIGURE 90B

203/240

TACGGCACCTGACCCAAAAACTTGATTAGGGTATGGTCACGTAGTGGGCCATCGC
 CCTGATAGACGGTTTCGCCCTTGACGTTGGAGTCCACGTTCTTAATAGTGGACTCT
 TGTTCAAACTGGAACAACACTCAACCTATCTCGGTCTATTCTTTGATTATAAGGGA
 TTTTGCCTGATTTGCCCTATTGGTAAAAAATGAGCTGATTTAACAAATATTAACCGA
 ATTTAACAAAATATTAACGTTACAATTTCGCCCTGATGCGGTATTTCTCCTAACGCAT
 CTGTCGCGTATTCACACCGCATACCGGATCTGCGCAGCACCATGGCTGAAATAACCT
 CTGAAAGAGGAACCTGGTAGGTTACCTCTGAGGCGGAAAGAACCGAGCTGTGGAATGTGT
 GTCAGTTAGGTGTGAAAGTCCCCAGGCTCCCAGCAGGCAGAAGTATGCAAAGCATGC
 ATCTCAATTAGTCAGCAACCCAGGTGAAAGTCCCCAGGCTCCCAGCAGGCAGAAGTA
 TGCAAAGCATGCACTCAATTAGTCAGCAACCATAGTCCCCCTAATCTCCGCCATCC
 CGCCCTAACTCCGCCAGTCCGCCATTCTCCGCCATGGCTGACTAATTTC
 TTATGCAAGAGGCCAGGCCCTCGGCTCTGAGCTATTCCAGAAGTAGTGTGAGGAGGCT
 TTTTGGAGGCCTAGGCTTTGCAAAAGCTGATTCTCTGACACAAACAGTCTGAACT
 TAAGACCATGGCCAAGCCTTGTCTCAAGAAGAATCCACCCCTCATTGAAAGAGCAACGGC
 TACAATCAACAGCATCCCCATCTGAAAGACTACAGCGTCCAGCGCAGCTCTCTAG
 CGACGGCCGATCTTCACTGGTGTCAATGTATATCATTTCAGTGGGACAGCCGACGGCAGT
 ACTCGTGGTGTGGCACTGCTGCTGCGGAGCTGGCAACCTGACTTGTATCGTGC
 GATCGGAAATGAGAACAGGGCATCTTGAGCCCTGCGACGGTCCGACAGGTGCTTCT
 CGATCTGCATCTGGGATCAAAGCCATAGTGAAGGACAGTGTGATGGACAGCCGACGGCAGT
 TGGGATTGCTGAATTGCTGCCCTGGTTATGTTGTGGGAGGGCTAACGACTTCGTGGCG
 AGTTGAAATGACCGACCAACGACGCCAACCTGCCATCACGATGGCGCAATAAAATA
 TCTTATTTCTTACATCTGTGTTGGTTTTGTGATCGATAGCGATAAGGATC
 CGCGTATGGTCATCTCAGTACAATCTGCTGCTGCCATAGTTAACGCCAGCCCGA
 CACCCGCCAACACCCGCTGACGCCCTGACGGGCTGCTGCTGCCATCCGCTTAC
 AGACAAGCTGTGACCGTCTCGGGAGCTGATGTGTCAGAGGTTTCACCGTACACCG
 AACCGCGAGACAAAGGGCTGTGATACGCTATTATAGGTTATGTCATGATA
 ATAATGGTTCTTAGACGTAGGTGGACTTTGGGAAATGTGCCGGAACCCCTATT
 TGTTTATTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCTGATAA
 ATGCTCAATAATATTGAAAAGGAAGAGTATGAGTATTCAACATTCCGTGTCGCCCTT
 ATTCCCTTTTGCGGATTTGCCCTCCTGTTTGCTACCCAGAAACGCTGGTAAA
 GTAAAAGATGCTGAGATCAGTGGTGCACGAGTGGTTACATCGAACTGGATCTCAAC
 AGCGGTAAGATCCTTGAGAGTTTGGGGGAAGAACGTTTCAATGATGAGCACTTT
 AAAGTTCTGCTATGTGGCGGGTATTATCCGTATTGACGCCGGCAAGAGCAACTCGGT
 CGCCGCATAACTATTCTCAGAATGACTGGTGTAGTACTCACAGTCACAGAAAAGCAT
 CTTACGGATGGCATGACAGTAAGAGAATTATGCACTGCTGCCATAACCAGTGTGATAAC
 ACTGCGGCCACTTACTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTG
 CACAACATGGGGATCATGTAACTCGCCCTGATGTTGGAACCGAGCTGAATGAAGCC
 ATACCAAACGACGAGCGTGAACACCACGATGCCCTGATGCAATGGCAACACGTTGCGAAA
 CTATTAACTGGCGAACTACTTAACCTAGCTTCCCGCAACAAATTAAAGACTGGATGGAG
 GCGGATAAAAGTTGCAGGACCACTTCTGCGCTGGCCCTCCGGTGGTTATTGCT
 GATAAACTGGAGCCGGTGGCGTGTGAGCTGTTACCGTACAGCAGCACTGGGCCAGAT
 GGTAAGCCCTCCGTATGTTACACGACGGGAGTCAGGCAACTATGGATGAA
 CGAAATAGACAGATCGCTGAGATAGGTGCCTACTGATTAAGCATTGGTAACTGTCAGAC
 CAAGTTACTCATATACTTAGATTGATTAAACTTCATTAAATTAAAGGATC
 TAGGTGAAGATCCTTTGATAATCTCATGACCAAATCCCTAACGTGAGTTCTGTT
 CACTGAGCGTCAGACCCGTAGAAAAGATCAAAGGATCTTCTGAGATCTTTCTG
 CGCGTAATCTGCTGCTGCAAACAAAAACCGCTACCGAGCGGTGGTTGCTGCC
 GATCAAGAGCTACCAACTCTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCA
 AATACTGTCCTCTAGTGTAGCGTAGTTAGGCCACCACTCAAGAAACTCTGAGCACCG
 CCTACATACCTCGCTCTGCTAATCTGTTACCACTGGCTGCTGCCAGTGGCGATAAGTCG
 TGTCTTACGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGCTGA
 ACGGGGGGTTCTGTCACACAGCCAGCTGGAGCGAACCTACACCGAAGTGTGAGATAC
 CTACAGCGTGAGCATTGAGAAAAGCGCCACGCTTCCGAAGGGAGAAAGGCGGACAGGTAT
 CCGGTAAGCGGAGGGTGGAAACAGGAGAGCGCACGAGGGAGCTTCAGGGGAAACGCC
 TGGTATCTTATAGTCCTGCGGTTGCGCACCTCTGACTTGAGCGTGTGATTGGTGA
 TGCTCGTCAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACCGGGCTTTACGGTTC
 CTGGCCTTTGCTGGCCTTGCTCACATGTTCTGCTGTTATCCCCTGATTCTGTTG
 GATAACCGTATTACCGCCTTGAGTGTGAGCTGATACCGCTGCCGAGCGAACGACCGAG-

FIGURE 90C

204/260

CGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCC
GCGCGTTGGCCGATTCATTAATGCAGAGCTTGCATTCGCGCTTTCAATATTATTGA
AGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAAT
AAACAAAATAGGGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACC
ATTATTATCATGACATTAACCTATAAAAATAGGCGTAGTACGAGGCCCTTCACTCATTA
G

FIGURE 9D

205/240

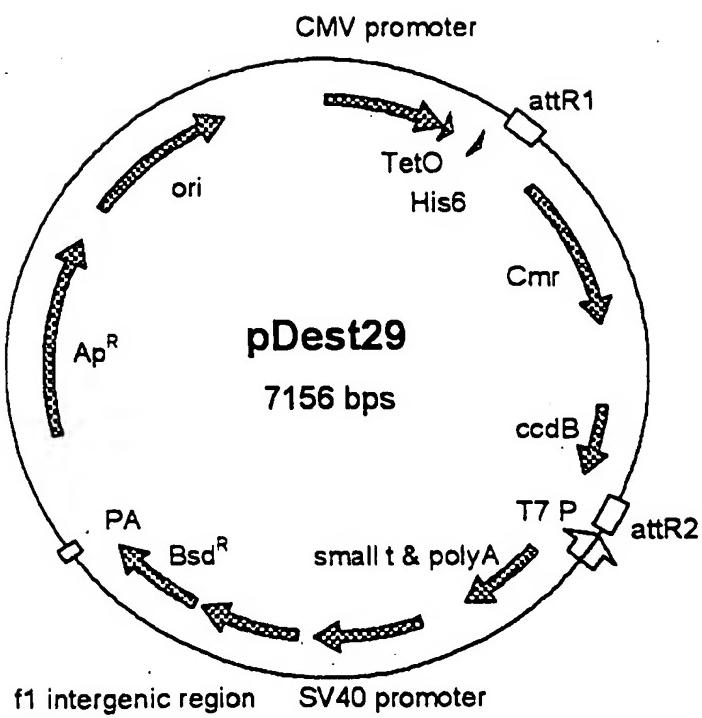


FIGURE 91 A

pDEST29 7156 bp

ATGCATGTCGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCGCCAACGACCCC
 CGCCCATGACGTCATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCAT
 TGACGTCATGGTGGAGTATTACGGTAAACTGCCCCTGGCAGTACATCAAGTGTAT
 CATATGCCAAGTACGCCCTATTGACGTCATGACGGTAAATGGCCGCTGGCATTAT
 GCCCAGTACATGACCTTATGGACTTCCACTTGGCAGTACATCAACGGTATTAGTCATC
 GCTATTACCATGGTGGATGCGGTTTGCAGTACATCAATGGCGTGGATAGCGGTTGAC
 TCACGGGAATTCCAAGTCTCCACCCATTGACGTCATGGAGTTGTCACCAA
 AATCAACGGACTTCCAAAATGCGTAACAACCGCCATTGACGCAAATGGCGGT
 AGGCGTGTACGGTGGAGGTCTATATAAGCAGAGCTCCCTATCAGTGATAGAGATCTC
 CCTATCAGTGATAGAGATCGTCAGCAGCTGTTAGTGAACCGTCAGATGCCCTGGAGA
 CGCCATCCACGCTGTTGACCTCAGAAGACACCGGACCGATCCAGCCTCCGGACC
 ATGGCGTACTACCATCACCATCACACGGTGGATATCCTCGAGGCCATACAAGT
 TTGTACAAAAAAAGCTGAACGAGAAACGTAAGATATAAAATATCAATATATTAAATTAG
 ATTTTGCAAAAAACAGACTACATAACTGTAACACAAACATATCCAGTCACTATGG
 CGGCCGATTAGGCACCCAGGCTTACACTTATGCTCCGGCTCGTATAATGTGTGGA
 TTTGAGTTAGGATCCGGCAGATTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAA
 TCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTGAGGCAT
 TTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTAGCTGGATATTACGGCCTTT
 TAAAGACCGTAAAGAAAAATAAGCACAAGTTATCCGGCTTATTACATTCTGCC
 GCCTGATGCAATGCTCATCCGGATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGGATAT
 GGGATAGTGGTACCCCTGTTACACCGTTTCCATGAGCAAACGTTTGTGGATATTACGGCCTTT
 TCTGGAGTGAATACCACGACGATTCCGGCAGTTCTACACATATACTCGCAAGATGTGG
 CGTGTACGGTAAAACCTGGCTATTCCCTAAAGGTTATTGAGAATATGTTTCC
 TCTCAGCCAATCCCTGGGTGAGTTTACCGAGTTTGTGATTTAACGTTGCAATATGGACA
 ACTTCTCGCCCCCGTTTACCATGGCAAATATTACGCAAGGGGACAAGGTGCTGA
 TGCCGCTGGGATTCAGGTTCATCATGCCGCTGTGATGGCTTCCATGCGCAGAACATGC
 TTAATGAATTACAACAGTACTGCGATGGCAGGGCGGGCGTAAACGCGTGGATCCG
 GCTTACTAAAAGCCAGATAACAGTATGCGTATTGCGCGCTGATTTGCGGTATAAGAA
 TATATACTGATATGTATACCGAAGTATGTCAAAAGAGGTGTGCTATGAAGCAGCGTAT
 TACAGTGACAGTTGACAGCGACAGCTATCAGTGCTCAAGGCATATATGATGTCATATC
 TCCGGTCTGGTAAGCACAACCATGCGAATGAAAGCCGTCGTCTGCGTGGCAACGCTGG
 AAAGCGAAAATCAGGAAGGGATGGCTGAGGTGCGCCGGTTATTGAAATGAAACGGCTCT
 TTTGCTGACGAGAACAGGGACTGGTGGAAATGCAAGTTAACGGTTACACCTATAAAAGAGA
 GAGCCGTTATCGTCTGTTGAGTACAGAGTGTGATATTGACACGCCCCGGCGACG
 GATGGTGATCCCCCTGGCCAGTGCACGTCTGCTGAGATAAGCTCCCGTGAACCTTA
 CCCGGTGGTGCATATCGGGGATGAAAGCTGGCGCATGATGACCAACCGATATGGCCAGTGT
 GCGGCTCCGTTATCGGGGAGAAGTGGCTGATCTCAGCCACCGGAAAATGACATCAA
 AAACGCACTAACCTGATGTTCTGGGAATATAATGTCAGGCTCCGTTATACACAGCCA
 GTCTGCAAGGTCGACCATAGTGACTGGATATGTTGTTACAGTATTATGTTAGTCTGTT
 TTTTATGCAAAATCTAATTAAATATGATATTATCATTTACGTTCTCGTTCA
 CTTTCTGTACAAAGTGGTGGCTGAGGGCCGCTCTAGAGGGCCAAGCTTACCGTGCAT
 GCGACGTCATAGCTCTCCCTATAGTGAGTGTGATTATAAGCTAGGCACTGGCGTGT
 TTACAAACGTCGTGACTGGAAAAGCTGCTAGCTGGATTTGAGGTTACAGTAAAGCTCTA
 CTGTTGAGCATAATTGACAAACTACCTACAGAGATTAAAGCTCTAAGGTTAAAT
 AAAATTAAAGTGTATAATGTTAAACTAGCTGCATATGCTTGTGCTTGGAGAGTTT
 GCTTACTGAGTATGATTATGAAATATTACACAGGAGCTAGTGATTCTAATTGTTG
 TGTATTGAGATTACAGTCCCAAGGCTCATTTCAGGCCCTCAGTCCTCACAGTCTGTT
 CATGATCATAATCAGCCATACACATTGAGGTTTACTGCTTAAAAAACCTCCC
 ACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTAACTTGTGTT
 TGCACTGCTTAAATGGTACAAATAAGCAATAGCATTACAAATTCACAAATAAGCATT
 TTTTCAGTCATTCTAGTTGTTGCTGAGGCTTACAGTCAATGATCTTATCATGTC
 GATCGATCCTGCAATTAGAATGCCAACGCGCGGGAGAGGGCGTTGCGTATTGGCT
 GCGTAATAGCGAAGAGGCCGACCGATGCCCTCCAAACAGTGTGCGCAGCCTGAATG
 GCGAATGGGACGCGCCCTGAGCGCGCATTAGCGCGGGGTGTTGCGTACGCGCA
 GCGTACCGCTACACTTGCAGCGCCCTAGGCCGCTCCTTCGCTTCTCCCTTC
 TTCTGCCACGTTGCCGCTTCCCCGTCAAGCTCTAAATCGGGGCTCCCTTAGGGT-

FIGURE 91B

207/240

TCCGATTTAGTGCTTACGGCACCTCGACCCCCAAAAACTTGATTAGGGTATGGTAC
 GTAGTGGGCCATCGCCCTGATAGACGGTTTCGCCCCCTTGACGTTGGAGTCCACGTTCT
 TTAATAGTGGACTCTGTTCCAAACTGGAACAACACTCAACCCCTATCTGGTCTATTCTT
 TTGATTTATAAGGGATTTGCCGATTCGGCCTATTGGTTAAAAAATGAGCTGATTTAAC
 AAATATTAACCGAATTTAACAAAATATTAACTGTTACAAATTTCGCTGATGCCGTAT
 TTTCTCCTACGCATCTGTGCGGTATTACACCGCATACGGATCTGCCAGCAGCACCAT
 GGCCTGAAATAACCTCTGAAAGAGGAACCTGGTAGGTTACCTCTGAGGCCGAAAGAAC
 AGCTGTGGAATGTGTCAGTTAGGGTAGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAA
 GTATGCAAAGCATGCATCTCAATTAGTCAGCAACCCAGGTGAGAAAGTCCCCAGGCTCCC
 CAGCAGGAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCGCC
 TAACTCCGCCATCCGCCCTAACCTCGCCAGTCCGCCATTCTCCGCCATGGCT
 GACTAATTTTTATTATGCAAGAGGCCGAGGCCCTGGCTCTGAGCTATTCCAGA
 AGTAGTGAGGAGGCTTTTGAGGCCAGGCTTTGCAAAAAGCTTGATTCTCTGACA
 CAACAGTCGAACCTAACGACATGGCAAGCCTTGCTCAAGAAGAATCCACCCAT
 TGAAAGAGCAACGGCTACAATCAACAGCATCCCCATCTGAGACTACAGCGTCGCCAG
 CGCAGCTCTCTAGCGACGCCGATCTCACTGGTAGTCAATGTATATCATTTACTGG
 GGGACCTGTGAGAAGCTGTGGTGCTGGCACTGCTGCTGCCAGCTGGCAACCT
 GACTTGATCGCGATCGAAATGAGAACAGGGCATCTTGAGCCCTGCCAGGGTG
 CGCACAGGTGCTCTGATCTGCACTGGATCAAAGCCATAGTGAAGGACAGTGTGG
 ACAGCCGACGGCAGTTGGATTGTAATTGCTGCCCTTGTTATGTGTGGAGGGCTA
 AGCACTTCGTGGCCAGTTGAAATGACCGACCAAGCGACGCCAACCTGCCATCAGAT
 GGCGCAATAAAATATCTTATTTCATTACATCTGTGTTGGTTTTGTGTGAATCG
 ATAGCGATAAGGATCCCGTATGGTCACTCTCAGTACAATCTGCTCTGATGCCGATAG
 TTAAGCCAGCCCCACACCCGCAACACCCGCTACGCCCTGACGGGCTTGTCTGCTC
 CGGCATCCGTTACAGACAAGCTGTGACCGCTCCGGAGCTGCATGTGTAGAGGTT
 TCACCGTCATACCGAAACCGCGAGACGAAAGGGCTCGTACGCCATTTCACCCA
 GTTAATGTCATGATAATAATGGTTCTTAGACGTCAGGTGCACTTTCGGGAAATGTG
 CGCGGAACCCCTATTGTTATTCTAAATACATTCAAATATGTATCGCTCATGAGA
 CAATAACCCCTGATAAAATGCTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACAT
 TTCCGTGTCGCCCTTATTCCCTTTGCCGATTTGCCCTCTGTTTGTCAACCCA
 GAAACGCTGGTAAAGTAAAGATGCTGAAGATCAGTTGGTGACGAGTGGTTACATC
 GAACTGGATCTCAACAGCGTAAGATCCTTGAGAGTTTCCGCCGAAGAACGTTTCCA
 ATGATGAGCACTTTAAAGTTCTGCTATGTGGCGGGTATTATCCGTATTGACGCCGG
 CAAGAGCAACTCGTCGCCGATACACTATTCTAGAATGACTTGGTTGAGTACTCACCA
 GTCACAGAAAAGCATCTACGGATGGCATGACAGTAAGAGAATTATGCAGTGTGCCATA
 ACCATGAGTGATAACACTGCCCAACTTACTCTGACAAACGATCGAGGACCGAAGGAG
 CTAACCGCTTTTGACAAACATGGGGATCATGTAACTCGCCTGATCGTTGGGAACCG
 GAGCTGAATGAAGCCATACCAAACGACGAGCGTACACCAAGATGCCGTAGCAATGGCA
 ACAACGTTGCACAAACTATTAACTGGCAACTACTACTCTAGTCCCGGCAACAATT
 ATAGACTGGATGGAGGCGGATAAAAGTGTGCAAGGACCACTCTGCGCTGCCCTCCGGCT
 GGCTGGTTATTGCTGATAAAATCTGGAGGCCGGTGGCTGGGTCTCGCGGTATATTGCA
 GCACTGGGCCAGATGGTAAGCCCTCCGTATGTTAGTTACACGACGGGAGTCAG
 GCAACTATGGATGAACGAAATAGACAGATCGTGGAGATAGGTGCTACTGATTAAGCAT
 TGGTAACTGTCAGACCAAGTTACTCATATATACTTTAGATTGATTAACCTCATGAC
 AAATTTAAAGGATCTAGGTGAAGATCCTTTGATAATCTGACCAAAATCCCTAA
 CGTGAGTTCTGTCACGCGCTAGACCCCTGAGAAAAGATCAAAGGATCTTCTGA
 GATCCTTTCTGCGCTAATCTGCTGCTGAAACAAAAACCCGCTACCGCG
 GTGGTTGTTGCCGGATCAAGAGCTACCAACTCTTCCGAAGGTAACGGCTTCAGC
 AGAGCGCAGATACCAAATACTGTCCTCTAGTGAGCCGTTAGGACCCACTTCAAG
 AACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCTGTTACCAAGTGGCTGCG
 AGTGGCGATAAGTGTGTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
 CAGCGGTGGCTGAACGGGGGGTCTGTCACACAGCCAGCTGGAGCGAACGACCTAC
 ACCGAACGAGATACCTACAGCGTGGAGGAGGAGCGCACGAGGGAGCTT
 AAGGCCGACAGGTATCGGTAAGCGGAGGGTGGAAACAGGAGAGCGCACGAGGGAGCTT
 CCAGGGGAAACGCCCTGGTATCTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAG
 CGTCGATTTGATGCTGTCAGGGGGCGAGCCTATGAAAAACGCCAGCAACGCG
 GCCTTTTACGGTCTGCCCTTGCTGGCTTTGCTCACATGTTCTCTGCGTTA
 TCCCCTGATTCTGAGTAAACCGTATTACCGCTTGAGTGAGCTGATACCGCTGCCGC-

FIGURE 91c

208/240

AGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCCAATACGC
AAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGAGCTTGCATTGCCTCGCTT
TTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATAACATATTGAA
TGTATTAGAAAAATAAACAAATAGGGGTTCCCGCGCACATTCCCCGAAAAGTGCCACCT
GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTAGTACGAGG
CCCTTCACTCATTAG

FIGURE 91D

209/240

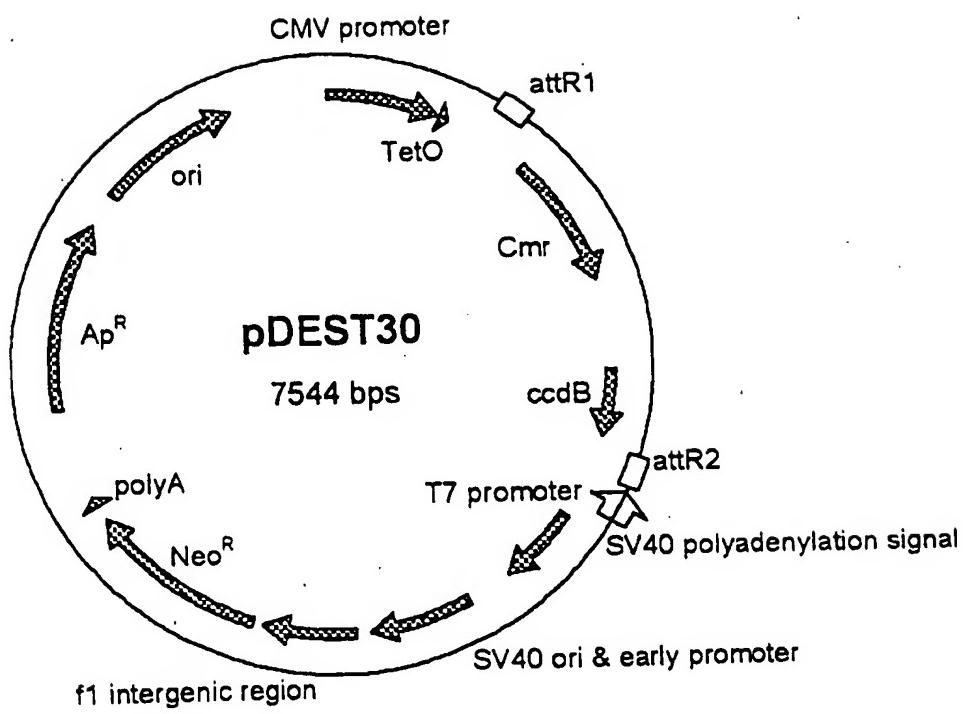


FIGURE 92A

pDEST30 7544 bp

ATGCATGCGTTACATAACTACGGTAAATGGCCGCCTGGCTGACGCCCAACGACCCC
 CGCCCATGACGTCAATAATGACGTATGTCCTAGTAACGCCAATAGGA^{CTT}CCAT
 TGACGTCAATGGTGGAGTATTACGGTAAACTGCCACTGGCAGTACATCAAGTGTAT
 CATATGCCAAGTACGCCCTATTGACGTCAATGACGGTAAATGGCCGCTGGCATTAT
 GCCCAGTACATGACCTTATGGACTTCTACTTGGCAGTACATCTACGTATTAGTCATC
 GCTATTACCATGGTGATGCGGTTTGGCAGTACATCAATGGCGTGGATAGCGGTTGAC
 TCACGGGATTCCAAGTCTCCACCCATTGACGTCAATGGAGTTGTTGGCACCAA
 AATCAACGGACTTCCAAAATGCGTAACAACCTCCGCCATTGACGAAATGGCGGT
 AGGCGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCAGTGTAGAGATCTC
 CCTATCAGTGTAGAGATCGCAGCTCGTTAGTGAACCGTCAGATCGCTGGAGA
 CGCCATCCACGCTTTGACCTCATAGAACAGACCCGGACCGATCCAGCCTCCGGACT
 CTAGAGGATCCCTACCGGTGATATCCTGAGGCCATCAACAAGTTGTCACAAAAAGCTG
 AACGAGAAAACGTAATGATATAATCAATATATTAAATTAGATTGTCATAAAAAAC
 AGACTACATAATACTGTAAAACACAACATATCCAGTCACTATGGCGCCGATTAGGCAC
 CCCAGGCTTACACTTATGCTTCCGGCTCGTATAATGTGCGATTGAGTTAGGATCC
 GGCGAGATTTCAAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGATATACCAC
 CGTTGATATATCCAATGGCATCGTAAAGAACATTGAGGCAATTTCAGTCAGTTGCTCA
 ATGTACCTATAACCAGACCGTTCAAGTGGATATTACGCCCTTTAAAGACCGTAAAGAA
 AAATAAGCACAAGTTTATCCGGCTTATTCAATTCTGGCCCTGATGAATGCTCA
 TCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTATATGGGATAGTGTTCACCC
 TTGTTACACCGTTTCCATGAGCAAACCTGAAACGTTTCACTCGCTGGAGTGAATACCA
 CGACGATTCCGGCAGTTCTACACATATATTGCAAGATGTGGCTGTTACGGTGAAA
 CCTGGCCTATTCCCTAAAGGGTTATTGAGAATATGTTTCTGTCAGGCCAATCCCTG
 GGTGAGTTCAACAGTTTGATTTAACAGTGGCAATATGGACAACCTTCTCCCCCCGT
 TTTCACCATGGGCAAATATTATACGCAAGGCAGAACAGGTGCTGATGCCGCTGGCGATTCA
 GGTCATCATGCCGCTGTGATGGCTCCATGTCGGCAGATGCTTAATGAATTACAACA
 GTACTGCGATGAGTGGCAGGGGGGGCTAAAGATCTGGATCCGGTTACTAAAGCCAG
 ATAACAGTATGCGTATTGCGCTGATTTGCGGTATAAGAATATATACTGATATGTA
 TACCCGAAGTATGTCAAAAGAGGTGCTGATGAAGCAGCGTATTACAGTGAAGTGTAC
 AGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAGTAAATCTCCGGTCTGGTAAGCA
 CAACCATGCGAGATGAAGCCGTCGTCAGCGTGGAAAGCGGAAAATCAGG
 AAGGGATGGCTGAGGTCGCCGGTTATTGAAATGAACGGCTTTGCTGACGAGAAC
 GGGACTGGTAAATGCGATTAAAGGTTACACTATAAAAGAGAGAGCCGTTATCGTCTG
 TTGTTGGAATGTCAGAGTGTGATATTGACACGCCGGCGACGGATGGTATCCCCCTG
 GCCAGTCACGCTGCTGTCAGATAAGTCTCCGTGAACTTACCCGGTGGTGCATATC
 GGGGATGAAAGCTGGCGCATGATGACCAACCGGATATGCCAGTGTGCCGGTCTCCGTTATC
 GGGGAAGAAGTGGCTGATCTGCCACCGGAAATGACATCAAAACGCCATTAAACCTG
 ATGTTCTGGGAATATAATGTCAGGCTCCCTATACACAGCCAGTCTGCAGGTGACCA
 TAGTGAATGAGTATGTTGTTACAGTATTATGATGCTGTTTATGCAAATCTA
 ATTAAATATATTGATATTATCATTACGTTCTCGTTCACTGTTACAAAGT
 GGTTGATGGCGGCCCTAGAGGGCCAAGCTACGCGTGCATGCCAGTCATAGCTC
 TCTCCCTATAGTGAATGCTGATTATAAGCTAGGCAGTGGCGTGTGTTACACGTCGTGA
 CTGGGAAACTGCTAGCTGGATCTTGTGAAAGGAACCTACTTCTGTTGACATA
 ATTGGACAAACTACAGAGATTAAAGCTCAAGGAAATATAAAATTAAAGTGT
 ATAATGTTAAACTAGCTGCATATGCTGCTGAGAGTTGCTACTGAGTATGA
 TTTATGAAAATATTACACAGGAGCTAGTGTGATTCTAATTGTTGTTAGATTCA
 CAGTCACGCTCAGTGGCTTACAGTCTGTCAGTGTGTTACATGATCATAATCAG
 CCATACCACATTGTAGAGGTTTACTTGTGTTAAAAACCTCCACACCTCCCCCTGAA
 CCTGAAACATAAAATGAATGCAATTGTTGTTAACTGTTATTGCAAGCTATAATGG
 TTACAAATAAGCAATAGCATCACAATTCAAAATAAGCATTCTTCACTGCATT
 TAGTTGTTGTTGTCACAGTCTGTCAGTGTGTTACATGCTGGATCGATCCTGCA
 AATGAATCGGCCAACGCGGGAGAGGCGGTTGCGTATTGGCTGGCTAATAGCGAAG
 AGGCCCGACCGATGCCCTCCCAACAGTTGCGCAGCCTGAATGGCAATGGACGCC
 CCTGTAGCGGGCATTAGCGCGGGTGTGGTGTGTTACGCGCAGCGTACAC
 TTGCCAGGCCCTAGGCCCGCTCTTCGCTTTCTCCCTTCGACCGTAC
 CGGCTTCCCGTCAAGCTAAATCGGGGCTCCCTTAAAGGTTCCGATTAGTGTCTT-

FIGURE 92B

211/240

TACGGCACCTCGACCCCCAAAAAACTGATTAGGGTATGGTTACGTAGTGGGCCATCGC
 CCTGATAGACGGTTTCGCCCTTGACGTTGGAGTCCACGTTCTTAATAGTGGACTCT
 TGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTGATTTAAGGGAA
 TTTGCCGATTCGGCTATTGGTAAAAAATGAGCTGATTAACAAATATTTAACCGCA
 ATTTAACAAAATATTAACGTTACAATTTCGCCTGATGCGGTATTTCTCCTACGCAT
 CTGTCGGTATTCACACCGCATACCGGATCTGCGCAGCACCATGCCCTGAAATAACCT
 CTGAAAGAGGAACCTGGTAGGTACCTCTGAGGCGGAAAGAACCCAGCTGGAATGTGT
 GTCAGTTAGGGTGTGGAAAGTCCCAGGCTCCCAAGCAGGAGAAGTATGCAAAGCATGC
 ATCTCAATTAGTCAGCAACCAAGGTGTGGAAAGTCCCAGGCTCCCAAGCAGGAGAAGTA
 TGCAAAGCATGCAATTAGTCAGCAACCATAGTCCCAGGCTCCCAAGGCTCCCAAGCAGGAGAAGTA
 CGCCCTAACTCCGCCAGTCCGCCATTCTCCGCCATGGCTGACTAATTTTTA
 TTATGCAGAGGCCGAGGCCCTCGGCTCTGAGCTATTCCAGAAGTAGTGGAGGAGGCT
 TTTGGAGGCCTAGGCTTTGCAAAAGCTGATTCTCTGACACAAACAGTCTCGAACT
 TAAGGCTAGAGCCACCATGATTGAAACAGATGGATTGACCGAGGTTCTCCGGCGCTTG
 GGTGGAGAGGCTATCGGCTATGACTGGGCAACACAGACAATCGGCTGCTGATGCCGC
 CGTGTCCGGCTGTGCGCAGGGCGCCGGTCTTTTGTCAGAAGACCGACCTGTCCGG
 TGCCCTGAATGAATGCACTGCAAGGAGGAGCAGCGGCTATCGTGGCTGCCACGACGGCGT
 TCCTTGCAGCTGCTGCCAGTGTCACTGAAGCGGGAGGGACTGGCTGCTATTGGG
 CGAAGTCCGGGGCAGGATCTCTGTCATCTCACCTGCTCCTGCCAGAAAGTATCCAT
 CATGGCTGATGCAATGCGCGGCTGCACTACGCTGATCGGCTACCTGCCATTGACCA
 CCAAGCGAAAATCGCATCGAGCGAGCACGTAACCGATGGAAGCGGGCTTGTGCGATCA
 GGATGATCTGGACGAAGAGCATCAGGGCTCGGCCAGCGAACACTGTCGCCAGGCTCAA
 GGCACATGCCGACGGCAGGATCTGTCGTGACCCATGGCGATGCGCTTGTGCCGAA
 TATCATGGTGGAAATGGCCGCTTTCTGGATTGACATGACTGTGGCCGGCTGGGTGTGGC
 GGACCGCTATCAGGACATAGCAGCTGGTACCCGTGATATTGCTGAAGAGCTTGGCGGCGA
 ATGGGCTGACCGCTCTCGTCTTACGGTATCGCCGCTCCGATTGCAAGCGCATCGC
 CTTCTATCGCCTCTTGACGAGTTCTGAGCGGGACTCTGGGTTGAAATGACCGAC
 CAAGCGACGCCAACCTGCCATCACGATGCCAATAAAATATCTTATTTCATTACA
 TCTGTGTGGTTGGTTGGTATAGGTTAATGTCATGATAAAGATCCGCTATGGTCACTCT
 CAGTACAATCTGCTCTGATGCCCATAGTTAAGCCAGCCCCGACACCGCAAACACCGC
 TGACGCCCTGACGGCTGTCTGCTCCGGATCCGCTTACAGACAAGCTGTGACCGT
 CTCCGGAGCTGCACTGTCAGAGGTTTCAACCGTACCGAAACCGCGAGACGAAA
 GGGCCTCGTGATACGCTATTAGGTTAATGTCATGATAAATGGTTCTTAGAC
 GTCAGGGCACTTTCGGGAAATGCGCGGAACCCCTATTGTTATTCTAAAT
 ACATTCAAATATGATCCGCTCATGAGACAATAACCCGTATAATGCTTCAATAATATTG
 AAAAGGAAGAGTATGAGTATTCAACATTCCGTGCGCCTTATTCCCTTTGCGGC
 ATTGCGCTTCTGTTGTCACCCAGAAACGCTGGTAAAGTAAAGATGCTGAAGA
 TCAGTTGGTGCACTGGTACATGCAACTGGATCTAACAGCGTAAGATCCTTGA
 GAGTTTCGCCCGAAGAACGTTTCCAATGATGAGCACTTTAAAGTTCTGCTATGTGG
 CGCGGTATTATCCGTATTGACGCCGGCAAGAGCAACTCGTCGCCATACACTATTG
 TCAGAATGACTGGTGTGAGTAACGACAGAAAGCATTACGGATGGCATGAC
 AGTAAGAGAATTATGCACTGCTGCCATAACCATGAGTATAACACTCGGCCAACTTACT
 TCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTGCAACAACATGGGGATCA
 TGTAACTCGCCTTGATCGTGGGACCGGAGCTGAATGAGACCCATACAAACGAGCG
 TGACACCACGATGCCGTAGCAATGCCAACACGTTGCCAAACTATTAACTGGCAACT
 ACTTACTCTAGCTTCCCGCAACAATTAAAGACTGGATGGAGGGGATAAAAGTTGCA
 ACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTATTGCTGATAAATCTGGAGCCGG
 TGAGCGTGGGTCTCGCGGTATCATTGCACTGGGGCAGATGGTAAGGCCCTCCGTAT
 CGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAAACGAAATAGACAGATCGC
 TGAGATAGGTGCTCACTGATTAAGCATTGGTAACGTCAAGACAAAGTTACTCATATA
 ACTTAGATTGATTAAAACCTCATTTAATTAAAGGATCTAGGTGAAGATCCTTT
 TGATAATCTCATGACCAAAATCCCTAACGTTGAGTTCTGTTCACTGAGCGTCAGACCC
 CGTAGAAAAGATCAAAGGATCTTCTGAGATCTTCTGCGCTGCGTAATCTGCTGCTT
 GCAAACAAAAAACACCGCTACCGGGTGGTTGTCGCCAGATACCAAAACTGTCTTCTAGT
 TCTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAAACTGTCTTCTAGT
 GTAGCCGTAGTTAGGCCACCACTTCAGAAGACTCTGTCAGCACCGCTACATACCTCGCTCT
 GCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCATAAGTCGTCTTACCGGGTTGGA
 CTCAAGACGATAGTTACCGATAAGGCCAGCGGCTGACGGGGTTCGTGCAC-

FIGURE 92C

212/240

ACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATA CCTACAGCGTGAGCATTG
AGAAAGCGCCACGCTTCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGT
CGGAACAGGGAGAGCGCACGAGGGAGCTTCAGGGGAAACGCCTGGTATCTTTATAGTCC
TGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTGTGATGCTCGTCAGGGGGCG
GAGCCTATGGAAAAACGCCAGCAACGCCCTTTTACGGTTCTGGCCTTTGCTGGCC
TTTGCTCACATGTTCTTCCTCGGTATCCCCGTATTCTGTGGATAACCGTATTACCGC
CTTGAGTGAGCTGATACCGCTCGCCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAG
CGAGGAAGCGGAAGAGCGCCAATCGCAAACGCCCTCTCCCGCGCGTTGGCCGATTCA
TTAATGCAGAGCTGCAATTGGCGCTTTCAATATTATTGAAGCATTTATCAGGGTTA
TTGTCTCATGAGCGGATACATATTGAATGTATTAGAAAAATAAACAAATAGGGGTTCC
GCGCACATTCGGAAAAGTGCCACCTGACGCTTAAGAAACCATTATTATCATGACATT
AACCTATAAAATAGCGTAGTACGAGGCCCTTCACTCATTAG

FIGURE 92D

213/2ho

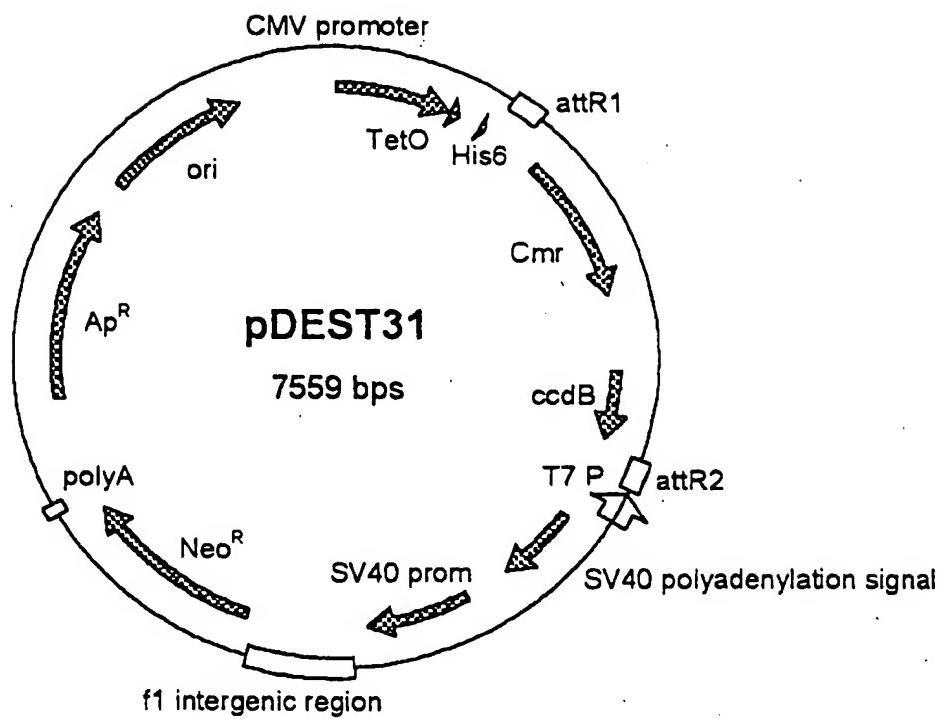


FIGURE 93A

pDEST31 7559 bp

ATGCATGTCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCC
 CGCCCATTGACGTCATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCAT
 TGACGTCATGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTAT
 CATATGCCAAGTACGCCCTATTGACGTCATGACGGTAAATGGCCCGCCTGGCATTAT
 GCCCAGTACATGACCTTATGGACTTCCACTTGGCAGTACATCTACGTATTAGTCATC
 GCTATTACCATGGTGATGCCGTTTGGCAGTACATCAATGGCGTGGATAGCGGTTGAC
 TCACGGGATTTCCAAGTCTCCACCCATTGACGTCATGGAGTTGTTGGCACCAA
 AATCAACGGGACTTCCAAAATGTCGTAACAACCTCGCCCAATTGACGCAAATGGCGGT
 AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCCCTATCAGTGTAGAGATCTC
 CCTATCACTGATAGAGATCGTCGACGAGCTGTTAGTGAACCGTCAGATCGCCTGGAGA
 CGCCATCCACGCTTTGACCTCCATAGAACACCCGGGACCGATCCAGCCTCCGGACC
 ATGGCGTACTACCACCATCACCACACCCGGTGTATCCTCGAGGCCATCACAAGT
 TTGTACAAAAAAAGCTGAACGAGAAACGTTAAATGATATAAATATCAATATTTAAATTAG
 ATTTGCTATAAAAACAGACTACATAACTGTAACACAAACATATCCAGTCACTATGG
 CGGCCGATTAGGCACCCAGGTTTACACTTATGCTTCCGGCTGTATAATGTGTGGA
 TTTGAGTTAGGATCCGGCAGATTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAA
 TCACTGGATATACCACCGTGTATATCCCAATGGCATCGTAAAGAACATTGAGGCAT
 TTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTAGCTGGATATTACGGCTTT
 TAAAGACCGTAAAGAAAAATAAGCACAAGTTTATCCGGCTTATTCACATTCTGCC
 GCCTGATGAATGTCATCCGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGTAT
 GGGATAGTGTTCACCCCTGTTACCCGTTTACCGTGTATCCATGAGCAAACCTGAAACGTTTATCGC
 TCTGGAGTGAATACCACGACGATTCCGGCAGTTCTACACATATATTGCAAGATGTGG
 CGTGTACGGTAAAACCTGGCTATTCCCTAAAGGGTTTATTGAGAATATGTTTTCG
 TCTCAGCCAATCCTGGGTGAGTTTACCACTGGCAAATATTACGCAAGGCGACAAGGTGCTGA
 ACTTCTCGCCCCGTTTACCATGGCAAATATTACGCAAGGCGACAAGGTGCTGA
 TGCGCTGGCGATTAGGTTCATCATGCCGCTGTGTAGGCTTCCATGCGGAGAATGC
 TTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGCGTAAACCGTGGATCCG
 GCTTACTAAAAGCCAGATAACAGTATGCGTATTGCGCGTGATTTTGCCTGATAAGAA
 TATATACTGATATGTATACCGAAGTATGTCAAAAGAGGTGTGCTATGAAGCAGCGTAT
 TACAGTGACAGTTGACAGCGACAGCTACGTTGCTCAAGGATATGATGTCAATATC
 TCCGGTCTGGTAAGCACAACCATGCGAAATGAGCCGCTGCTGCGTGCAGAATGCTGG
 AAAGCGAAAATCAGGAAGGGATGGCTGAGGTCGCCGGTTATTGAAATGAACGGCTCT
 TTTGCTGAGGAAACAGGGACTGGTGAATGCAAGTTAAGGTTACACCTATAAAAGAGA
 GAGCGTTATCGTCTGTTGGATGTACAGAGTGTATATTGACACGCCGGCGACG
 GATGGTGTACCCCTGGCAGTGCACGTCTGCTGAGATAAGTCTCCGTGAACTTAA
 CCCGGTGGTCATATCGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCAGTGT
 GCCGGTCTCGTTATCGGGAAAGAAGTGGCTGATCTCAGCCACCGGAAAATGACATCAA
 AAACGCCATTAACCTGATGTTCTGGGAATATAATGTCAGGCTCCGTTATAACAGCCA
 GTCTGAGGTGCGACCATAGTGAATGGATATTGTTGTTTACAGTATTATGTTAGTCTGTT
 TTTATGCAAAATCTAATTAAATATTGATATTATCATTTACGTTCTCGTTCAG
 CTTTCTGTACAAAGTGGTGTAGGGCGCCGCTTAGAGGGCCCAAGCTTACGCGTGCA
 GCGACGTCAAGCTCTCCCTATAGTGAGTCGTATTATAAGCTAGGCACGGCCGCTCGT
 TTTACACGTCGTGACTGGAAAATGCTAGCTGGATCTTGTGAAGGAACCTTACTT
 CTGTTGTCGACATAATTGGACAAACTACCTACAGAGATTAAAGCTCTAAGGTAATAT
 AAAATTAAAGTGTATAATGTGTTAAACTAGCTGCAATGCTGCTGCTGAGAGTTT
 GCTTACTGAGTGTGTTATGAAAATATTACACAGGAGCTAGTGATTCTAATTGTTG
 TGTATTAGATTACAGTCCCAAGGCTCATTCAGGCCCTCAGTCCTCACAGTCTGTT
 CATGATCATATACTGACCCATACCACTTGTAGAGGTTTACTTGTCTTAAAAACCTCCC
 ACACCTCCCCCTGAAACCTGAAACATAAAATGAATGCAATTGTTGTTAACTTGT
 TGCACTTATAATGGTTACAATAAGCAATAGCATCACAAATTCAACAAATAAAGCATT
 TTTTCACTGCAATTCTAGTTGTGGTTGTCCAAACTCATCAATGTATCTTATCATGTC
 GATCGATCCTGCAATTGAATGCCAACGCCGGAGAGGCCGGTTGCGTATTGGCT
 GGCAGTAAAGCGAAGAGGCCGACCGATGCCCTCCCAACAGTTGCGCAGCCTGAATG
 GCGAATGGGACGCCCTGTAGCGCGCATTAAAGCGCCGGGTGTGGTGGTTACGCCA
 GCGTACCGCTACACTGCCAGGCCCTAGGCCCGCTCTTCCGTTCTCCCT
 TTCTGCCACGTTGCCGGCTTCCCGTCAAGCTCTAAATGGGGCTCCCTTGTAGGGT-

FIGURE 93B

TCCGATTTAGTCCTTACGGCACCTCGACCCCCAAAAACTGATTAGGGTGTGGTCAC
GTAGTGGGCCATCGCCCTGATAGACGGTTTCGCCCTTGACGTTGGAGTCACGTTCT
TTAATAGTGGACTCTGTTCAAACCTGGAACAAACACTCAACCCTATCTCGGTCTATTCTT
TTGATTATAAGGGATTTGCCATTGCCCTATTGGTAAAAATGAGCTGATTTAAC
AAATTTAACCGGAATTAAACAAATATTACGTTACAATTGCCGTGCGGTAT
TTCTCCTAACGCATCTGCGGTATTACACCGCATACCGGATCTGCGCAGCACC
GCCCTGAAATAACCTCTGAAAGAGGAACCTGGTAGGTACCTCTGAGGGGAAAGAAC
AGCTGTGGATGTGTCAAGTTAGGGTGTGAAAGTCCCAGGCTCCCCAGCAGGCAGAA
GTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGAAAGTCCCAGGCTCCC
CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGTCCCAGGCT
TAACCTCGCCCCATCCGCCCTAACCTCGCCAGTCCGCCATTCTCGCCCCATGGCT
GACTAAATTTTTATTTATGAGGGCCGAGGCCGCTGGCCTCTGAGCTATTCCAGA
AGTAGTGAGGAGGCTTTTGAGGCCCTAGGCTTGTGAAAAGCTGATTCTCTGACA
AACAGTCTCGAACTTAAGGCTAGGCCACCATGATTGAAACAAGATGGATTGACGCAGG
TTCTCGGGCGCTTGGGTGGAGAGGCTATTGGCTATGACTGGGACAACAGACAATCGG
CTGCTCTGATGCCCGTGTCCGGCTGAGCGCAGGGCGCCGGTCTTTGTCAA
GACCGACTGTCCGGTGCCTGAAATGAACTGCAGGACGAGGAGCGCGGCTATGTGGCT
GCCACGACGGCGTCTCTGCGAGCTGTGCTGACGTTGCACTGAAGCGGGAAAGGGA
CTGGCTGCTATTGGCGAAGTGCCTGGAGGATCTCTGTCATCTCACCTGCTCCTGC
CGAGAAAGTATCCATCATGGCTGATGCAATGCGCGGCTGCATACGCTGATCCGGCTAC
CTGCCATTGACCAAGCAGAACATGCATCGAGCGAGCACCTACTCGATGGAAGC
CGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGCTCGGCCAGCGAACT
GTTGCCAGGCTCAAGGCCGATGCCGACGGCGAGGATCTGTCGTGACCCATGGCGA
TGCCGCTTGCCGAAATATCATGGTGGAAATGGCGCTTTCTGATTATCGACTGTGG
CCGGCTGGGTGTGGCGAACCGCTATCAGGACATAGCGTTGGCTACCGTGATATTGCTGA
AGAGCTTGGCGCGAATGGCTGACCGCTTCTCGTGTCTTACGGTATGCCGCTCCCGA
TTCGCAAGCGATGCCCTCTATGCCCTCTTGACGAGTTCTCTGAGCGGGACTCTGGGG
TCGAATGACCGACCAAGCAGGCCAACCTGCCATACGATGCCGAAATAAAATATC
TTTATTTCATACATCTGTTGGTTTTGTGTGAATCGATAGCGATAAGGATCCG
CGTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGATAGTTAAGCCAGCCCGACA
CCCGCCAACACCCGCTGACGCCCTGACGGGCTGTGCTGCCGATCCGCTTACAG
ACAAGCTGTGACCGTCTCGGGAGCTGCATGTCAGAGGTTTACCGTACACCGAA
ACGCCGAGACGAAAGGGCTCGTACGCCTATTATAGGTTAATGTCATGATAAT
AATGGTTCTAGACGTAGGTGGACTTTGGGAAATGTGCGCGAACCCCTATTG
TTTATTTCTAAATACATTCAAATATGATCCGCTCATGAGACAATAACCTGATAAAAT
GCTTCATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTCCGTTGCGCCCTAT
TCCCTTTTGCGGATTTCGCCCTCTGTTGCTACCCAGAAACGCTGGTAAAGT
AAAAGATGCTGAAGATCAGTGGTGCACGAGTGGTTACATCGAACTGGATCTAACAG
CGTAAGATCCTGAGAGTTTGCCCCGAAAGAACGTTTCAATGATGAGCACTTTAA
AGTTCTGCTATGTGGCGCGGTATTATCCGTATTGACGCCGGCAAGAGCAACTCGGTG
CCGCATACACTATTCTCAGAATGACTGGTTGAGTACTCACAGTCACAGAAAAGCATCT
TACGGATGGCATGACAGTAAGAGAATTATGCACTGCTGCTGCCATAACCATGAGTGATAAACAC
TGCGCCAACTTACTCTGACAACGATCGGAGGACGAAAGGAGCTAACCGTTTTGCA
AACATGGGGATCATGAACTCGCTTGATGTTGGAACCGGAGCTGAATGAAGCCAT
ACCAACGACGAGCGTACACCGATGCCGTAGCAATGGCAACACGTTGCGCAAAC
ATTAACCTGGCGAACTACTTACTCTAGCTTCCCGCAACAATTAAAGACTGGATGGAGGC
GGATAAAAGTTGCAGGCCACTTCTGCGCTGGCCCTTCCGGCTGGCTGGTTATTGCTGA
TAAATCTGGAGGCCGGTGAAGCTGGTCTCGCGTATCATTGAGCACTGGGCCAGATGG
TAAGCCCTCCGTATCGTAGTTATCTACAGGACGGGAGTCAGGCAACTATGGATGAACG
AAATAGACAGATCGCTGAGATAGGTGCTCACTGATTAGGTAAGCATTGTAAGGACCA
AGTTTACTCATATATACTTTAGATTGATTAAAATTCATTAAATTTAAAGGATCTA
GGTGAAGATCCTTTGATAATCTCATGACCAAATCCCTAACGTGAGTTCTGTTCCA
CTGAGCGTCAGACCCGTAGAAAAGATCAAAGGATCTCTGAGATCCTTTTCTGCG
CGTAATCTGCTGCTGCAAACAAAAAACCCACCGCTACAGCGGTGGTTTGTGCGG
TCAAGAGCTACCAACTCTTTCCGAAAGGTAACGGCTCAGCAGAGCGCAGATAACCAA
TACTGCTCTAGTGTAGCCGTAGTTAGGCCACCTCAAGAACACTCTGAGCAGGCC
TACATACCTCGCTCTGCTAATCCTGTTACCAAGTGGCTGCTGCCAGTGGGAGTAAGTCGTG
TCTTACGGGGTGGACTCAAGACGGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAAC

FIGURE 93C

216/240

GGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATAACCT
ACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCCGACAGGTATCC
GGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACCGAGGGAGCTCCAGGGGAAACGCCTG
GTATCTTATAGTCCTGTCGGTTTGCACCTCTGACTTGAGCGTCGATTTTGATG
CTCGTCAGGGGGCGGAGCCTATGGAAAACGCCAGCAACGCCCTTTACGGTTCC
GGCCTTTGCTGGCCTTTGCTCACATGTTCTTCTGCGTTATCCCTGATTCTGTGGA
TAACCGTATTACCGCCTTGAGTGAGCTGATACCGCTGCCGCAGCGAACGACCGAGCG
CAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGGCCCAATACGCAAACGCCCTCCCCGC
GCGTTGGCGATTCAATGCAAGAGCTTGCATTGGCTTCAATATTATTGAAG
CATTTATCAGGGTTATTGTCATGAGCGGATACATATTGAATGTATTTAGAAAAATAA
ACAAATAGGGGTTCCGCGCACATTTCCCCGAAAGTGCCACCTGACGTCTAAGAAACCAT
TATTATCATGACATTAACCTATAAAATAGGCAGTAGTACGAGGCCCTTCACTCATTAG

FIGURE 93D

217/240

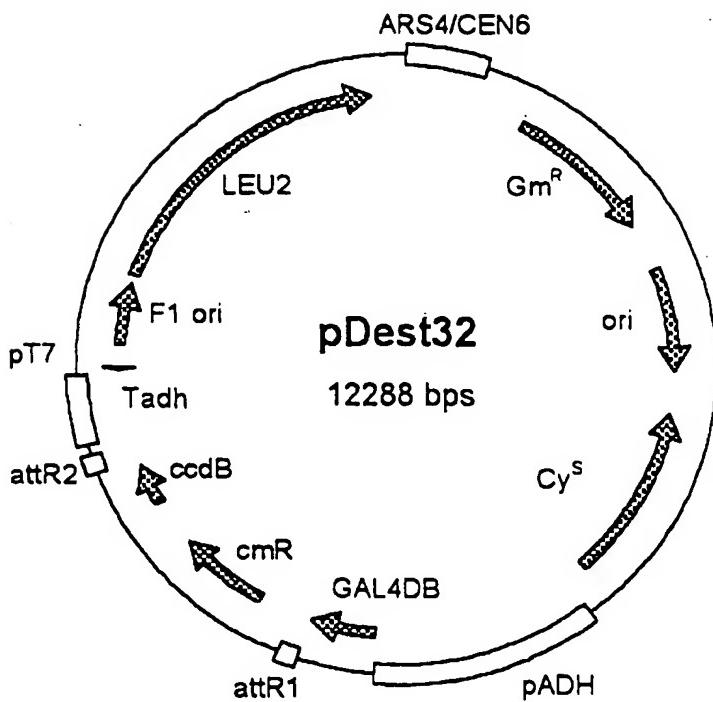


FIGURE 94A

218/240

pDEST32 12288 bp

GACGAAAGGGCCTCGTATACGCCTATTTTATAGGTTAATGTCATGATAATAATGGTT
 CTTAGGACGGATCGCTGCCTGTAACTTACACGCGCTCGTATCTTTAATGATGGAATA
 ATTTGGGAATTACTCTGTGTTATTTATTTATGTTTGATTTGGATTTAGAAAAGT
 AAATAAAAGAAGGTAGAAGAGTTACGGAATGAAGAAAAAAATAAACAAAGGTTAAAAAA
 ATTTCAACAAAAAGCGTACTTTACATATATATTATTAGACAAGAAAAGCAGATTAATA
 GATATACATTGATTAACGATAAGTAAAATGTAATCACAGGATTTCGTGTGGTCT
 TCTACACAGACAAGATGAAACAATTGGCATTATACTGAGAGCAGGAAGAGCAAGATA
 AAAGGTAGTATTGTTGGCGATCCCCCTAGAGTCTTTACATCTTCGGAAAACAAAAACT
 ATTTTTCTTAATTCTTTTACTTTCTATTAAATTATATATTATTTATTTAATTTAATTTA
 ATTTAAATTATAATTATTTTATAGCACGTGATGAAAAGGACCCAGTGGCATTTCGG
 GGAAATGTGCGCGAACCCCTATTGTTATTTCTAAATACATTCAAATATGTATCCG
 CTCATGAGACAATAACCCCTGATAATGCTTCAATAATCTGAGTGGCAGGGCCGTGTC
 TCAAAATCTCTGTGTTACATTGACAAGATAAAAATATCATCATGAAACAATAAAACT
 GTCTGCTTACATAAACAGTAATAACAGGGGTGTTATGAGCCATATTCAACGGGAAACGTC
 TTGCTGGAGGCCGCGATTAAATTCCAACATGGATGCTGATTATATGGGTATAATGGC
 TCGTAGCCAACCACACTAGAACTATAGCTAGAGTCCTGGCGAACAAACGATGCTCGCCTT
 CCAGAAAACCGAGGATGCGAACCCACTTCATCCGGGTGAGCACCACGGCAAGCGCCGCG
 ACGGCCGAGGTCTCCGATCTCTGAAGCCAGGGCAGATCCGTGACAGCACCTTGCCTG
 AGAAGAACAGCAAGGCCCAATGCCGACGATGCGTGGAGACCGAAACCTTGCCTCGT
 TCGCCAGGCCAGGACAGAAATGCCGACTTCGCTGCTGCCAGGGTGCCTGAGC
 CACCGTGGAAACGGATGAAGGCACGAACCCAGTTGACATAAGCCTGTCGGTCTGTAAC
 TGTAATGCAAGTAGCGTATGCGCTCACGCAACTGGTCCAGAACCTTGAACGACGAGCG
 GTGGTAACGGCCAGTGGGGTTTCATGGCTGTTATGACTGTTTTTGACAGTCTA
 TGCCTGGGCATCCAAGCAGCAAGCCGTTACGCCGTGGGTCATGTTGATGTTATGGA
 GCAGCAACGATGTTACGCAAGCAGATGTTACGCAGCAGGGCAGTCGCCCTAAACAA
 AAGTTAGGTGGCTCAAGTATGGCATTCAGCACATGTAGGCTGGCCCTGACCAAGTC
 AAATCCATGCCGGCTGCTCTGATCTTCGGTCTGAGTTGGAGACGTAGCCACCTAC
 TCCCACATCAGCCGACTCCGATTACCTCGGGAACTTGCTCCGTAGTAAGACATTCA
 GCGCTTGCCTTCGACCAAGAAGCGGTTGGCGCTCTCGCGCTTACGTTCTGCC
 AGGTTGAGCAGCCGCTAGTGAGATCTATATCTATGATCTCGAGTCTCCGGCAGCAC
 CGGAGGCAGGGCATTGCCACCGCCTCATCAATCTCTCAAGCATGAGGCCAACCGC
 GGTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGAGTGGCTCTAT
 ACAAAAGTTGGCATAACGGGAAGAAGTGATGCACTTTGATATCGACCCAAGTACCG
 TAACAATTGTTCAAGCCGAGATCGGCTCCGGCTAATAGGTTGATTGATGTTGGAC
 GAGTCGAATCGCAGACCGATAACAGGATCTGCCATCTATGGAACCTGCTCGGTGAGT
 TTTCTCTTCAATTACAGAAACGGTTTTCAAAATATGGTATTGATAATCTGATATGA
 ATAAATTGCAAGTTCAATTGATGCTGATGAGTTTTCTAATCAGAAATTGGTTAATTGGT
 TGTAACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGNATGACCAAAATCC
 AACGTGAGTTCTGTTCACTGAGCGTCAGACCCGTAAGAAAAGATCAAAGGATCTT
 GAGATCCTTTCTGCGCTAATCTGCTGTTGCAACAAAAACCCACCGCTACCA
 CGGTGGTTGTTGCCGATCAAGAGCTACCAACTCTTCCGAAGGTAACCTGCTTCA
 GCAGAGCGCAGATACCAAATACTGCTCTAGTGTAGCCGTAGTTAGGCCACCACTTC
 AGAACTCTGAGCACCCTACATACCTCGCTCTGCTAATCTGTTACAGTGGCTGCTG
 CCAGTGGCGATAAGTCGTCCTACGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
 CGCAGCGGTGGGCTGAAAGGGGGTTGTCGTGACACAGCCCAGCTGGAGCGAACGAC
 ACACCGAAGTGGAGATACCTACAGCGTGGAGCATTGAGAAAGCGCCACGCTTCCGAAGGG
 GAAAGCGGACAGGTATCCGTAAGCGGAGGGTGGAAACAGGAGAGCGCACGAGGGAGC
 TTCCAGGGGAAACGCCCTGGTATCTTATAGTCTGTCGGGTTTCGCCACCTCTGACTTG
 AGCGTCGATTTGATGCTCGTACGGGGGCCGAGCCTATGGAAAACGCCAGCAACG
 CGGCCCTTTACGGTTCTGGCCTTTGCTGGCCTTGTGACTCACATGTTCTCTGCGT
 TATCCCCTGATTCTGTGATAACCGTATTACCGCCTTGAGTGAGCTGATACCGCTGCC
 GCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCA
 GCAAACCGCCTCTCCCGCGCTGGCCGATTCATTAAATGCACTGGCACGACAGGTTTC
 CCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAAATGAGTTACCTCACTCATTAG
 CACCCCAGGCTTACACTTATGCTTCCGGCTCTATGTTGAGTACCTCACTCATTAG
 AACAAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAGCTCGGAATTACCC
 -

FIGURE 94B

FIGURE 94C

TCAATATATTAAATTAGTTGCATAAAAAACAGACTACATAATACTGTAAAACACAAC
 ATATCCAGTCACTATGGCGGCCGCTAAGTGGCAGCATCACCGACGCACTTTGCGCCGA
 ATAAATACCTGTGACGGAAGATCACTTCGAGAATAAATAAATCCTGGTGTCCCTGTTGA
 TACCGGGAAAGCCCTGGGCCAACTTTGGCAAAATGAGACGTTGATCGGCACGTAAGAGG
 TTCCAACCTTCACCATAATGAAATAAGATCACTACCGGGCGTATTTTGAGTTATCGAG
 ATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGATATACCACCGTTGA
 TATATCCAATGGCATCGTAAAGAACATTGAGGCATTTCAGTCAGTTGCTCAATGTAC
 CTATAACCAGACCGTTAGCTGGATATTACGGCCTTTAAAGACCGTAAAGAAAAATAA
 GCACAAGTTTATCCGGCCTTATTACACATTCTGCCGCCTGATGAATGCTCATCCGGA
 ATTCCGTATGGCAATGAAAGACGGTGAGCTGGTATGGGATAGTGTTCACCCCTGTTA
 CACCGTTTCCATGAGCAAACCTGAAACGTTTCATCGCTCTGGAGTGAATACCAACGACGA
 TTTCCGGCAGTTCTACACATATAATTGCAAGATGTGGCGTGTACGGTGAACACCTGGC
 CTATTTCCCTAAAGGGTTATTGAGAATATGTTTCTGTCAGCCAATCCCTGGGTGAG
 TTTCACCACTTTGATTTAACGTTGCAATATGGACAACCTCTCGCCCCGTTTCA
 CATGGGAAATATTACGCAAGGCACAAGGGTGTGATGCCGCTGGCATTAGGTTCA
 TCATGCCGTCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTG
 CGATGAGTGGCAGGGCGGGCGTAATCTAGAGGATCCGGCTACTAAAGCCAGATAACA
 GTATGCGTATTTGCGCCTGATTTGCGGTATAAGAATATATACTGATATGTATACCCG
 AAGTATGTCAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTTGACAGCGAC
 AGCTATCAGTTGCTCAAGGCATATATGATGTCAATATCTCCGCTGGTAAGCACAACCA
 TGCAGAATGAAGCCCCTGCTGCGTGCCGAACGCTGGAAGCGGAAATCAGGAAGGGGA
 TGGCTGAGGTGCCCGTTATTGAAATGAACGGCTCTTGCTGACGAGAACAGGGACT
 GGTGAAATGCACTTAAGGTTACACCTATAAAAGAGAGAGAGCGTTATCGTCTGTTG
 GATGTCAGAGTGTATTATTGACACGCCCGGGCGACGGATGGTGTATCCCCCTGGCCAGT
 GCACGTCGCTGTCAGATAAGTCTCCGTTGAACTTTACCCGGTGGTGCATATCGGGGAT
 GAAAGCTGGCGCATGATGACCCACCGATATGGCACTGTGCGCGTCTCGTTATCGGGGAA
 GAAGTGGCTGATCTCAGCCACCGCGAAAATGACATCAAAACGCCATTAAACCTGATGTT
 TGGGAATATAAAATGTCAGGCTCCCTATACACAGCCAGTCTGCAGGTCGACCATAGTGA
 CTGGATATGTTGTTTACAGTATTATGTAGTCTGTTTATGCAAAATCTAATTAA
 TATATTGATAATTATCATTTACGTTCTCGTCAGCTTCTTGACAAAGTGGTTG
 ATGGCCGCTAAGTAAGTAAGACGCTCGAGCTAAGTAAGTAACGCCGCCACCGCGGTGG
 AGCTTGGACTTCTCGCCAGAGGTTGGTCAAGTCTCCAACTAACAGTTGTCGGCTTGT
 TACCTTGCCAGAAATTACGAAAAGATGGAAAAGGGTCAAATCGTTGGTAGATACGTTG
 TGACACTTCTAAATAAGCGAATTCTATGATTATGATTATTATTAATAAGTTAT
 AAAAAAAATAAGTGTATACAAATTAAAGTGTACTCTTAGGTTTAAACAGAAAATTCT
 GTTCTTGAGTAACTCTTCTGTAGGTTAGGTTCTCAGGTATAGCATGAGGTCGC
 TCTTATTGACCAACACCTCTACCGCATGCCAGCAATGCCAAATCGCTCCATT
 CACCCAAATTGAGATATGCTAACCTCAGCAATGAGTTGATGAATCTCGGTGTATT
 TGTCTCAGAGGACAATACCTGTTGATACGTTCTCCACACGGATCCAATTGCCCTA
 TAGTGTAGTCGTATTACAAATTCACTGGCGTCGTTTACAACGTCGTGACTGGAAAACC
 TGGCGTTACCCAACTTAATGCCCTGCAAGCACATCCCCCTTCCAGCTGGCGTAATAG
 CGAAGAGGCCGCACCGATGCCCTTCCAAACAGTTGCGAGCCCTGAATGGCGAATGGAC
 GCGCCCTGTAGCGCGCATTAAGCGGGCGGGTGTGGTGGTTACGCCAGCGTACCGCT
 ACACCTGCCAGGCCCTAGGCCCGCTCTTCTGCTTCTCCCTTCTGCCACG
 TTCGCCGGCTTCCCGTCAAGCTCTAAATCGGGGCTCCCTTAGGGTTCCGATTAGT
 GCTTACGGCACCTGACCCAAAAACTGATTAGGGTGTGGTACGTAAGTGGC
 TCGCCCTGATAGACGGTTTTCGCCCTTGACGTTGGAGTCCACGTTCTTAATAGTGA
 CTCTTGTCCAAACTGGAACAAACACTCAACCCATCTCGGTCTATTCTTTGATTATAA
 GGGATTGCGATTCCGCTATTGGTAAAAATGAGCTGATTAAACAAAATTAAAC
 CGAATTAAACAAATATAACGTTACAATTCTGATGCCGTATTTCTCCTACGC
 ATCTGTGCCGTATTCACACCGCATATGACCGGTCGAGGAGAACTCTAGTATATCCAC
 ATACCTAATATTATTGCTTATAAAAGGAACTGGAAACAATTACATCAAATCCACAT
 TCTCTCCTAAATCAATTGCTCTGTACTTCCTGTTCATGTTGTTCAAAACGTTATATT
 TATAGGATAATTATACTCTATTCTCAACAAGTAATTGGTTGTTGCCGAGCGGTCTAA
 GGCGCCTGATTCAAGAAATATCTGACCGCAGTTAACGTTGGGAATACTCAGGTATCGTA
 AGATGCAAGAGTTCGAATCTCTAGCAACCATTATTTTCTCAACATAACGAGAAC
 CACAGGGCGCTATCGCACAGAATCAAATTGATGACTGAAATTGGTTAATTTCAG
 AGGTCGCCCTGACCGATATACTTTCAACTGAAAATTGGAGAAAAGGAAAGGTGAG-

FIGURE 94D

AGGCCGGAACCGGCTTTCAATAGAATAGAGAACGCTCATGACTAAATGCTGCATCA
CAATACCTGAAGTGTACAATATTATTAAAGGACCTATTGTTTTCCAATAGGTGGTTAG
CAATCGTCTTACTTTCTAACCTTCTACCTTACATTCAAGCAATATATATATATT
TCAAGGATAACCATTCTAATGTCGCCCCATGCTGCCCTAAGAACATCGTCGTTT
GCCAGGTGACCACGGTGTCAAGAACATCACAGCGAACGCCATTAAGGTTAAAGCTAT
TTCTGATGTTGTTCCAATGTCAGGTCGATTCGAAAATCATTAAATTGGTGGTGCTGC
TATCGATGCTACAGGTGTCCACTTCAGATGAGGCGCTGGAAGGCCCAAGAACGGTTGA
TGCCGTTTGTAGGTGTGGTCTAAATGGGTACCGGTAGTGTAGACCTGA
ACAAGGTTTACTAAAATCCGTAAAGAACCTCAATTGTAAGGCCAACTTAAGACCATGTA
CTTGCACTCGACTCTTTAGACTATCTCCAATCAAGCCACAATTGCTAAAGGTAC
TGACTTCGTTGTTGTCAGAGAATTAGTGGGAGGTATTTACTTTGGTAAGAGAACAGGAAGA
CGATGGTGTGGTGCTGGGATAGTGAACAATACACCGTCCAGAACATGCAAAGAAT
CACAAGAATGGCCGTTTCATGGCCCTACAACATGAGGCCACCATGGCTATTGGTCTT
GGATAAAAGCTAATGTTTGGCCTCTCAAGATTATGGAGAAAATCTGTGGAGGAAACCAT
CAAGAACGAATCCCTACATTGAAGGTTCAACATCAATTGATTCTGCCGCCATGAT
CCTAGTTAAGAACCCACCACCTAAATGGTATTATAATCACCAGCAACATGTTGGTGA
TATCATCTCCGATGAAGCCCTCCGTTATCCCAGGTTCCCTGGGTTTGTGCCATCTGCGTC
CTTGGCCTCTTGCCAGACAAGAACACCGCATTGGTTGACCTATGCCACTATCTGTCTGCAAT
TGCTCCAGATTGCCAAGAATAAGGTTGACCTATGCCACTATCTGTCTGCAAT
GATGTTGAAATTGTCATTGAACTTGCTGAAGAACGGTAAAGGCCATTGAAGATGCAAGTTAA
AAAGGTTTGGATGCAAGGTATCAGAACTGGTAGTTAGGTGGTCCAACAGTACCAACCGA
AGTCGGTGTGCTGCCAGAACAGGTTAAGAAAATCCTGCTTAAAGATTCTCTTT
TTTATGATATTGTCATAAAACTTTATAATGAAATTCTATAATAGAAACGACACGAAATT
ACAAAATGGAATATGTTCAAGGGTAGACGAAACTATACGCAATCTACATACTTAT
CAAGAACGGGAAAAAGGAGGAGTAGTAAAGGAATACAGGTAAAGCAATTGATAACTATGGC
TCAACGGTGTAAAGGAAAAAGAATTGCACTTTAACATTAATATTGACAAGGAGGAGGGCAC
CACACAAAAAGTTAGGTGTAAACAGAAAATCATGAAACTACGATTCTAATTGATATTGG
AGGATTCTCTAAAAAAAAAAACACAACAAATAAAAACACTCAATGACCTGACCAT
TTGATGGAGTTAAGTCATAACCTTCTGAACCATTTCCATAATGGTAAAGTTCCCT
AAGAATTCTACTGTCAGAACGGCTTACGACGTAGTCGATATGGTCACCTCAGTA
CAATCTGCTCTGATGCCGATAGTTAAGCCAGCCCCGACACCCGCAACACCCGCTGACG
CGCCCTGACGGGCTTGTCTGCCCTGGCATCCGTTACAGACAAGCTGTGACCGTCTCCG
GGAGCTGCATGTGTCAAGGGTTTCAACCGTCAACCGAACACGCGCGA

FIGURE 94E

222/240

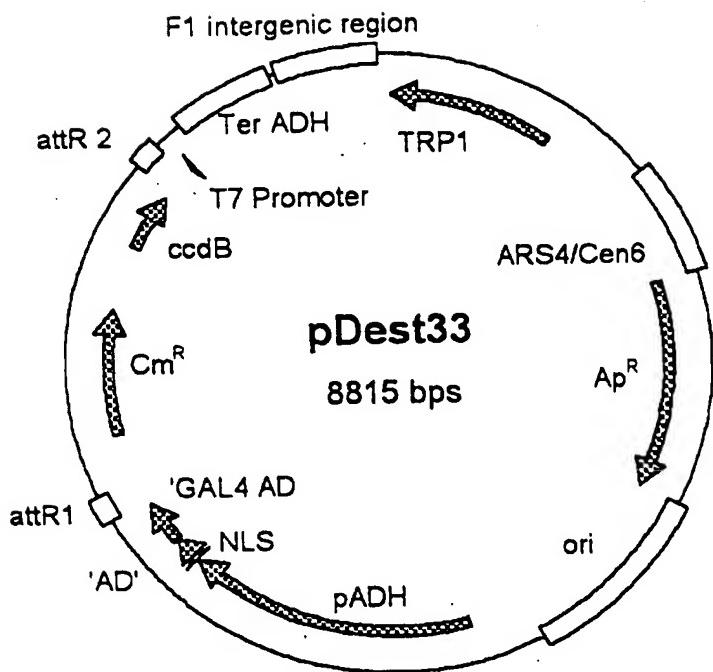


FIGURE 95A

223/240

pDEST33 8815 bp

GCCTTACGCATCTGTGCGGTATTCACACCGCAGGCAAGTGCACAAACAATACTTAAATA
 AATACTACTCAGTAATAACCTATTCTTAGCATTTGACGAAATTGCTATTGTTAG
 AGTCTTTACACCATTGTCACACCTCCGCTTACATCAAACCCAATAACGCCATTAA
 ATCTAAGCGCATACCAACATTCTGGCGTACGTCCACCAAGCTAACATAAAATGTAAGC
 TTTCGGGCTCTCTGCCTTCCAACCCAGTCAGAAATCGAGTTCAATCCAAAAGTTCAC
 CTGTCCCACCTGCTCTGAATCAAACAAGGGATAAACGAATGAGGTTCTGTGAAGCTG
 CACTGAGTAGTATGTTGCACTTGGAAATACGAGTCTTAAACTGGCAACCGA
 GGAACCTTGGAATTCTTGCCACGACTCATCTCCATGCAGTTGGACGATATCAATGCCGT
 AATCATTGACCAAGGCAAAACATCCCTTAGGTTACGAAACACGCCAACCAAGT
 ATTCGGAGTGCCTGAACTATTATGCTTTACAAGACTGAAATTTCCTGCAA
 TAACCGGGTCAATTGTTCTCTTCTATTGGGCACACATATAACCCAGCAAGTCAGCAT
 CGGAATCTAGAGCACATTCTGCGGCCCTGTGCTCTGCAAGCCGAAACTTCACCAATG
 GACCAAGACTACCTGTGAAATTAAATAACAGACATACTCCAAGCTGCCTTGTGCTTAA
 TCACGTATACTCACGTCTAACAGTCACTGGCCCTCTGGCCCTCTCTTTTC
 TTTTCGACCGAATTAAATTCTTAACTGGCAAAAAAGAAAAGCTCCGGATCAAGATTGT
 ACGTAAGGTGACAAGCTATTTCATAAAAGAATATCTTCAACTACTGCCATCTGGCGTC
 ATAACGCAAAGTACACATATATTACGATGCTGTCTATTAAATGCTCCTATATTATA
 TATAGTAATGTCGTTATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGATAGTTAA
 GCCAGCCCCGACACCCGCAACACCCGCTGACGCCCTGACGGGCTTGTCTGCCCG
 CATCCGTTACAGACAAGCTGTGACCGTCTCCGGAGCTGCATGTGTCAGAGGTTTAC
 CGTCATACCGAAACCGCGAGACGAAAGGGCTCGTGTACGCCATTAGGTTAA
 ATGTCATGATAATAATGGTTCTTAGGACGGATCGCTGCCGTAACTTACACGCCCTC
 GTATTTAAATGATGGAATAATTGGAATTACTCTGTTTATTAGTTATGTTT
 TGTATTGGATTAGAAAGTAAATAAGAAGGTAGAAGAGTTACGGAATGAAGAAAAAA
 AAATAACAAAGGTTAAAAAATTCAACAAAAAGCGTACTTACATATATTATTAG
 ACAAGAAAAGCAGATTAAATAGATATACATTGATTAAAGATAAGTAAATGAAATCA
 CAGGATTTCGTGTGGCTTCTACACAGACAAGATGAAACAATCGGATTAACCT
 GAGAGCAGGAAGAGCAAGATAAAAGGTAGTATTGTTGGCGATCCCCCTAGAGTCTTAA
 CATCTTCGAAACAAAAACTATTCTTAAATTCTTTTACTTCTATTAA
 TTTATATATTATTTAAATTATAATTATTAGCACGTGATGAAAG
 GACCCAGGTGGCACTTTCGGGAAATGTGCGCGAACCCCTATTGTTATTCTAA
 ATACATTCAAATATGTATCCGCTCATGAGACAATAACCTGATAATGCTTCAATAATAT
 TGAAAAGGAAGAGTATGAGTATTCAACATTCCGTGTCGCCCTATTCCCTTTGCG
 GCATTTCGCTTCTGTTGCTACCCAGAAACGCTGGTAAAGTAAAGATGCTGAA
 GATCAGTGGTGCACGAGTGGTTACATGAACTGGATCTCAACAGCGTAAGATCCTT
 GAGAGTTTCGCCCCGAAGAACGTTCCAATGATGAGCACTTTAAAGTTCTGCTATGT
 GGCGCGTATTATCCGTATTGACGCCGGCAAGAGCAACTCGGTGCCGCATACACTAT
 TCTCAGAATGACTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTACGGATGGCATG
 ACAGTAAGAGAATTATGCACTGCTGCCATAACCATGAGTGATAACACTGCCAACTTA
 CTTCTGACAACGATCGGAGGACGAAGGAGCTAACCGCTTTTCAACACATGGGGAT
 CATGTAACTCGCCCTGATGTTGGAAACCGGAGCTGAATGAAGCCATACAAACGACGAG
 CGTACACCCACGATGCCGTAGCAATGGCAACAAACGTTGCGCAAACATTAAACTGGCGAA
 CTACTTACTCTAGCTTCCCGCAACAAATTAAAGACTGGATGGAGGGCGATAAGTTGCA
 GGACCACTCTCGCCTCGGCCCTCCGGCTGGTTATTGCTGATAATCTGGAGCC
 GGTGAGCGTGGTCTCGCGGTATCTGCAAGCACTGGGGCCAGATGGTAAGGCCCTCCCGT
 ATCGTAGTTATCTACACGACGGGCAGTCAGGCAACTATGGATGAAACGAAATAAGACAGATC
 GCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTACTCATAT
 ATACTTATAGATTGTTAAACTCATTAAATTAAAGGATCTAGGTGAAGATCCTT
 TTTGATAATCTCATGACCAAATCCCTTAACGTGAGTTCTGTTCACTGAGCGTCAGAC
 CCCGTAGAAAAGATCAAAGGATCTTCTGAGATCCTTTTCTGCGCGTAATCTGCTGC
 TTGCAAACAAAAACACCGCTACCAAGCGGTGGTTGTTGCCGATCAAGAGCTACCA
 ACTCTTCCGAAGGTAACGGCTCAGCAGAGCGAGATAACCAAAACTGCTTCTA
 GTGTAGCCGTAGTTAGGCCACACTTCAAGAACACTGTCAGCACCCTACATACCTCGCT
 CTGCTAATCTGTTACAGTGGCTGCCAGTGGCATAAGTCGTGCTTACCGGGTTG
 GACTCAAGACGATAGTTACCGGATAAGGCCAGCGGTGGCTGAACGGGGGTTCGTGC
 ACACAGCCCAGCTGGAGCGAACGACCTACACCGAACACTGAGATAACCTACAGCGTGA
 GCGAT-

FIGURE 95B

TGAGAAAGGCCACGCTTCCGAAGGGAGAAAGGGACAGGTATCCGTAAGCGGCAGG
 GTCGGAACAGGAGAGCGCACAGGGAGCTTCAAGGGGGAACGCTGGTATCTTATAGT
 CCTGTCGGGTTGCCACCTCTGACTTGAGCGTCGATTTGTGATGCTCGTCAGGGGG
 CCGAGCTATGAAAAACGCCAGCAACGCCCTTTACGGTCTGGCCTTGTGG
 CCTTTGCTCACATGTTCTTCCGTTATCCCTGATTCTGTGGATAACCGTATTACC
 GCCTTGAGTGAAGCTGATACCGCTGCCGAGCCGAACGAGCGAGCGAGTCAGTG
 AGCGAGGAAGCGGAAGAGGCCAATACGCAAACGCCCTCCCCGCGCGTTGGCGATT
 CATTAAATGAGCTGGCACGACAGGTTCCGACTGGAAAGCGGGCAGTGACGCCAACGCA
 ATTAATGAGTTACCTCACTCATAGGCACCCAGGCTTACACTTTATGCTCCGGCT
 CCTATGTTGTGGAATTGTGAGCGGATAACAATTACACAGGAAACAGCTATGACCAT
 GATTACGCCAAGCTCGGAATTACCCCTACTAAAGGGAACAAAAGCTGGTACCGGGCCC
 CCCCTCGAGATCGGGATCGAAGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATG
 AAGGCAAAAGACAAATAAGGTCGAACGAAAATAAGTAAAAGTGTGATATGATG
 TATTGCTTGGCGCCGAAAAACGAGTTACGCAATTGACAATCATGCTGACTCT
 GTGGCGACCCCGCCTTGCGGCCCGGATAACGCTGGCGTGAGGCTGTGCCCGC
 GGAGTTTTGCGCTGCATTTCCAAGTTTACCCCTGCCATAAGGGCGAGATTGGAGA
 AGCAATAAGAATGCCGTTGGGTTGCGATGATGACGACCAACGACAATGGTGTCAATTAT
 TTAAGTTGCCGAAAGAACCTGAGTGCAATTGCAACATGAGTATACTAGAAGAAATGAGCCA
 AGACTTGCAGACCGAGTTGCCGTTGCGAACATAAGAGCAGCATGACCTTGAAG
 GTGAGACGCGCATAACGCTAGAGTACTTGAAGAGGAAACAGCAATAGGTTGCTACCA
 GTATAAATAGACAGGTACATACAACACTGAAATGGTTGCTGTTGAGTACGCTTCAA
 TTCATTGGGTGTGCACTTATTATGTTACAATATGGAAGGGAACTTACACTCTCTA
 TGCACATATAATTAAAGTCAATGCTAGTAGAGAAGGGGGTACACCCCTCCGCGC
 TCTTTCCGATTTCTAAACCGTGGATATTGCGATATCCTTTGTTGCTCCGG
 TGTACAATATGGACTTCCCTTTCTGCAACCAACCCATACTGGGATTCCCTATAAT
 ACCTTCGTTGGCTCCCTAACATGTTAGGCGAGGAGATAACAATAGAACAGATA
 CCAGACAAGACATAATGGGCTAAACAAGACTAACCAATTACACTGCTCATGGT
 GTACATAACGAACTAAACTGTTAGCCCTAGACTTGATAGCCATCATATCGAAGTTTC
 ACTACCTTTCCATTGCCATCTATTGAGTAATAATAGGCGCATGCAACTTTTC
 TTTTTTTCTCTCTCTCCCCGTTGTTGCTCACCATATGCCAATGACAAAAAA
 ATGATGGAAGACACTAAAGAAAAAAATAACGACAAAGACAGCACCAACAGATGCGTTG
 TCCAGAGCTGATGAGGGGTATCTCGAACACAGAAACTTTCTCTCCTTCATTCA
 AGCTATACCAAGCATACAATCAACTCAAGCTATGCCCAGAACAGGGAGGTCTCG
 AGCGCGCAATTAAATCAAAGTGGGATATTGCTGATAGCTCATTGCTCTTCA
 ACTAACAGTAGCAACGGTCCGAACCTCATAACAACACTCAAACAAATTCTCAAGCGCTTCA
 CAACCAATTGCCCTCTAACGTTCATGATAACTCATGAAATAAGAAATCACGGCTAGT
 AAAATTGATGATGGAATAATTCAAAACACTGTCACCTGGTTGGACGGACCAACTGCG
 TATAACCGTTGGAATCACTACAGGGATGTTAATACCAACTACAATGGATGATGTATAT
 AACTATCTATTGATGATGAAAGATAACCCACCAACCAAAAGAGGGTGGGTCGAAT
 CAAACAAAGTTGACAAAAAGCTGAACGAGAAACGTAAGGATATAATCAATATA
 TTAAATTAGATTGATGATGAAATAAGACTACATAATACTGTAACACAAACATATCCAG
 TCACTATGGCGGCCGCTAAGTTGGCAGCATCACCGACGCACCTTGCGCGAATAAAAC
 CTGTCAGGAAAGATCACTCGAGAATAAAATCTGGTGTCCCTGTTGATACCGGG
 AGCCCTGGGCCACTTGGCAGGGAAATGAGACGTTGATCGGCACGTAAGAGGTTCCA
 TTCACCATATAAGATCACTACCGGGCTATTGAGTTATGAGGATTTCA
 GAGCTAAGGAAGCTAAATGGAGAAAAAAACTGGATATACCAACCGTTGATATCCC
 AATGGCATGTAAGAACATTGAGGCATTCAGTCAGTGTCTCAATGTAACCTATAACC
 AGACCGTTAGCTGGATATTACGGCCTTTAAAGACCGTAAAGAAAAATAAGCACAAGT
 TTTATCCGGCCTTATTACACATTCTGCCCGCTGATGAAATGCTCATCCGGAAATTCCGTA
 TGGCAATGAAAGACGGTGAGCTGGTATGGGATAGTGTTCACCCCTGTTACACCGTT
 TCCATGAGCAAACGTTTCATCGCTCGGAGTGAATACCAACGACGATTCCGG
 AGTTTCTACACATATTGCAAGATGTTGCGTGTACGGTGAACACCTGGCCTATTCC
 CTAAAGGGTTATTGAGAATATGTTTCTCAGCCAATCCCTGGGTGAGTTTACCA
 GTTTGATTAAACGTGCCAATATGGACAACCTCTCGCCCCCGTTTACCATGGCA
 AATATTATACGCAAGGCAGAACAGGTGCTGATGCCGCTGGCGATTAGGTTCATGCG-

FILE 95C

TCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGGCATGAGT
GGCAGGGCGGGCGTAATCTAGAGGATCCGGCTACTAAAAGCCAGATAACAGTATGCCT
ATTGCGCGCTGATTTGCGGTATAAGAATATATACTGATATGTATACCCGAAGTATGT
CAAAAGAGGTGTCTATGAAGCAGCGTATTACAGTGACAGCGACAGCTATCA
GTTGCTCAAGGCATATATGAIGTCATACTCCGGTCTGGTAAAGCACAAACCATGCAGAAT
GAAGCCCGTCGTCTGCGCGAACGCTGGAAAGCGGAAATCAGGAAGGGATGGCTGAG
GTCGCCCGTTATTGAAATGAACGGCTTTGCTGACGAGAACAGGGACTGGTAAAT
GCAGTTAACGTTACACCTATAAAAGAGAGGCCGTTATCGTCTGGATGTACA
GAGTGATATTATTGACACGCCCGGGCGACGGATGGTATCCCCCTGGCCAGTCACGTCT
GCTGTCAGATAAAAGTCTCCGTGAACCTTACCCGGTGGTCATATCGGGATGAAAGCTG
GCGCATGATGACCACCGATATGCCAGTGTGCCGGTCTCGTTATCGGGAAAGAAGTGGC
TGATCTCAGCCACCGCGAAAATGACATCAAAACGCCATTACCTGATGTTCTGGGAAT
ATAAAATGTCAGGCCCGTTATACACAGCCAGTCAGGTGACCATAGTGACTGGATAT
GTTGTTTACAGTATTATGTTAGTCAGTTTATGCAAATCTAAATTAAATATATTGA
TATTATATCATTTCAGTTCTCGTTCAAGCTTCTGTACAAAGTGGTTGATGGCCGC
TAAGTAAGTAAGACGTCGAGCTCCCTATAGTGAGTCGATTAACACTGGCGTCGTTTAC
AACGCGTGAACGGGAAACACCGGTGAGCTAAGTAAGTAACGGCCGCCACCGCGGTG
GAGCTTGGACTCTCGCCAGAGGTTGGTCAGTCTCAAGGTTGTCGGCTTGT
CTACCTTGCAGAAATTACGAAAAGATGAAAGGGTCAAATCGTTGGTAGATACGTTG
TTGACACTTCTAAATAAGCGAATTCTTATGATTATGATTTTATTAAATAAGTTA
TAAAAAAATAAGTGATACAAATTAAAGTGACTCTTAGGTTTAAACGAAAATTCT
TGTCTTGAGTAACCTTCCGTAGGTCAAGGTTGCTTCAGGTATAGCATGAGGTGCG
CTCTTATTGACCACACCTCTACCGGCATGCCGAGCAAATGCCGTCAAATCGCTCCCCATT
TCACCCAATTGAGATATGCTAACTCCAGCAATGAGTTGATGAATCTCGGTGTGATT
ATGCTCAGAGGACAATACCTGTTGTAATCGTTCTCCACACGGATCCGCATCAGCGA
AATTGTAACGTTAATATTGTTAAATTGCGTTAAATATTGTTAAATCAGCTCATT
TTTAACCAATAGGCCGAAATGCCAAAATCCCTATAAAATCAAAGAATAGACCGAGAT
AGGGTTGAGTGTGTTCCAGTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAA
CGTCAAAGGGCGAAAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACCCTA
ATCAAGTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTAAAGGGAGCCC
CCGATTAGAGCTTGACGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAAGC
GAAAGGAGCGGGCGTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCAC
ACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTGCCATTCACTGCA

FIGURE 95D

226/240

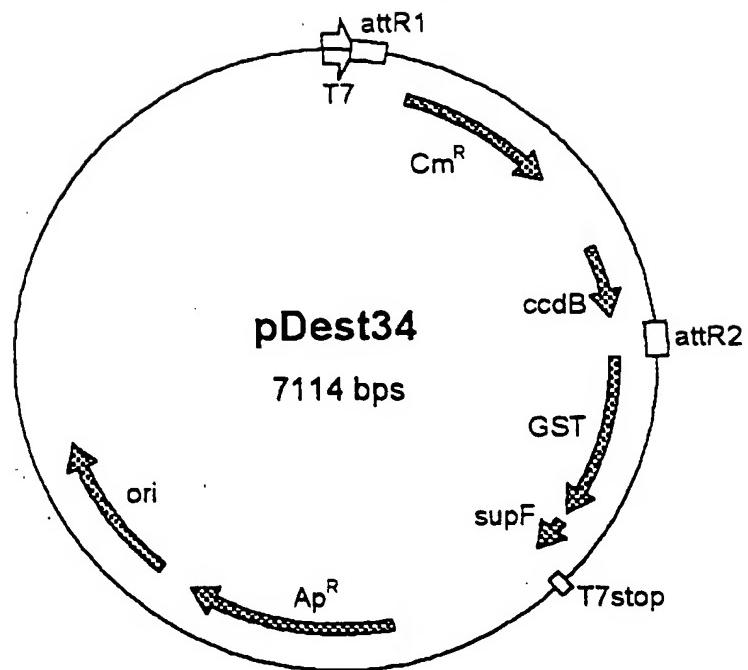


FIGURE 96A

227/240

pDEST34 7114 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
195..71	attR1
304..963	CmR
1305..1610	ccdB
1651..1775	attR2
1780..2472	GST
2675..2720	T7stop
3334..4194	ampR
4343..4982	ori

ATCGAGATCTCGATCCCGCAAATTAAACGACTCACTATAAGGGAGACCACAACGGTTTC
 CCTCTAGATCACAAGTTGTACAAAAAAGCTGAACGAGAACGTAAGTAAATGATATAAATAT
 CAATATATTAAATTAGATTTCGCATAAAAACAGACTACATAATACTGTAAACACAACA
 TATCCAGTCACTATGGCGGCCGCATTAGGCACCCCAGGCTTACACTTATGCTTCCGGC
 TCGTATAATGTGTGGATTGTAGTTAGGATCCGGCAGAGTTTCAGGAGCTAAGGAAGCT
 AAAATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAA
 GAACATTGAGGCATTCAGTCAGTGCTCAATGTACCTATAACCAAGACCGTTCAGCTG
 GATATTACGGCCTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTATCCGGCCTTT
 ATTACACATTCTGGCCGCTGATGAATGCTCATCCGAATTCCGTATGGCAATGAAAGAC
 GGTGAGCTGGTGATATGGGATAGTGTTCACCCCTGTTACCCGTTCCATGAGCAAAC
 GAAACGTTTCATCGCTCTGGAGTGAATACCAACGACGATTCCGGCAGTTCTACACATA
 TATTGCAAGATGTGGCGTTACGGTAAAAACCTGGCTATTCCCTAAAGGGTTATT
 GAGAATATGTTTCGTCTCAGCCAATCCCTGGGTGAGTTTACCAAGTGGTATTAAAC
 GTGGCCAATATGGACAACCTCTCGCCCCCGTTTACCCATGGGCAAATATTACGCAA
 GGCACAAAGGTGCTGATGCCGCTGGCGATTAGGTTCATCATGCCGCTGTGATGGCTTC
 CATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGCG
 TAAACCGTGGATCCGGCTACTAAAGCCAGATAACAGTATGCTATTGCGCGCTGAT
 TTTGCGGTATAAGAATATATACTGATATGTATACCGAAGTATGTCAAAAAGAGGTGTG
 CTATGAAGCAGCGTATTACAGTGACAGTGCAGCGACAGCTATCAGTGCTCAAGGCAT
 ATATGATGTCAATATCTCCGGCTGTGTAAGCACAACCATGCGAATGAAGCCGTCGTCT
 GCGTGGCGAACGCTGGAAAGCGGAAATCAGGAAGGGATGGCTGAGGTGCCCCGGTTAT
 TGAAATGAACGGCTTTGCTGACGAGAACAGGGACTGGTGAAATGCAAGTTAAGGTTT
 ACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTGTGATGTACAGGTGATATTATG
 ACACGCCGGCGACGGATGGTGAATCCCCCTGGCCAGTGCACGTGCTGTGATGACGAAAG
 TCTCCCGTGAACCTTACCCGGTGGTGCATATCGGGGATGAAAGCTGGCCATGATGACCA
 CCGATATGCCAGTGTGCCGGTCTCCGTTATCGGGGAAAGTGGCTGATCTCAGCCACC
 GCGAAATGACATCAAAACGCCATTAAACCTGATGTTGTTGGAAATATAAATGTCAGGCT
 CCCTTATACACAGCCAGTCTGCAGGTGCGACCATAGTGACTGGATATGTTGTTACAG
 TATTATGTAGTGTGTTTTATGCAAATCTAATTAAATATTGATATTGATTTATATCATT
 TACGTTCTCGTTCACTTCTGTACAAAGGGTGTGATATGTCCTTATGTTGTTGTTGTT
 TGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTTCTTGGATATTGATGTTGTTACAG
 TATGAAAGAGCATTGTATGAGCGCGATGAAGGGTGTAAATGGCAAAACAAAAGTTGAA
 TTGGGTTGGAGTTTCCAACTTCCCTTATTATATTGATGGTGTGTTAAATTAAACACAG
 TCTATGGGCATCATACGTATATAGCTGACAAGCACACATGTTGGGTGGTGTCAAA
 GAGCGTGCAGAGATTCAATGCTGAGGAGCGGTTGGATATTGATACGGTGTGTTCG
 AGAATTGCAATAGTAAAGACTTTGAAACTCTCAAAGTTGATTTCTTAGCAAGCTACCT
 GAAATGCTGAAAATGTTGAAAGATCGTTATGTCATAAAACATATTAAATGGTGTGATCAT
 GTAACCCATCCTGACTTCACTGTTGATGACGCTCTTGATGTTGTTTATACATGGACCCA
 ATGTGCCATGGATGCGTCCAAAATTAGTTGTTAAAAACGTATTGAAAGCTATCCCA
 CAAATTGATAAGTACTTGAAATCCAGCAAGTATAGCATGGCCTTGCAGGGCTGGCAA
 GCCACGTTGGTGGCGACCATCCTCCTAAATCGGATCTGGTCCCGCTCCATGGGGA
 TCCGGCTGCTAACAAAGCCCAGGGAAAGCTGAGTTGGCTGCTGCCACCGCTGAGCGCTT
 CCCGATAAGGGAGCAGGCCAGTAAAGCATTACCCGTGGTGGGGTTCCCGAGCGGGCAA
 GGGAGCAGACTCTAAATCTGCCGTATCGACTTCGAAGGTTGAAATCCTTCCCCCAC
 CATCACTTCAAAGTGAATTGCGTGTGAGCAATAACTAGCATAACCCCTGGGGCTCTAA-

FIGURE 96B

ACGGGTCTTGAGGGTTTGCTGAAAGGAGGAACATATCCGGATATCCACAGGACGG
 GTGTGGTCGCCATGATCGTAGTCGATAGTGGCTCCAAGTAGCGAAGCGAGCAGGACTG
 GGCAGCGGCCAAAGCGGTCGGACAGTGCCTCGAGAACGGGTCGCATAGAAATTGCATCA
 ACGCATATAGCGTAGCAGCACGCCATAGTGACTGGCAGTGTGGAATGGACGATAT
 CCCGCAAGAGGCCCGCAGTACCGCATAACCAAGCCTATGCCACAGCATCCAGGGTGA
 CGGTGCCAGGATGACGATGAGCGATGTTAGATTCATACCGTGCCTGACTGCGTT
 AGCAATTAACTGTGATAAAACTACCGCATTAAAGCTTATCGATGATAAGCTGTCAAACAT
 GAGAATTCTGAGAGACGAAAGGGCCTCGTGAACGGCTATTAGTTAGGTTAATGTCATG
 ATAATAATGGTTCTAGACGTAGGTGGCACTTTCGGGAAATGTGCGCGAACCCCT
 ATTTGTTATTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCCTGA
 TAAATGCTCAATAATATTGAAAAGGAGATGAGTATTCAACATTCCGTGTCGCC
 CTTATTCCCTTTGCGGCATTTCGCTTGTGACCCAGAACGCTGGTGA
 AAAGTAAAAGATGCTGAAGATCAGTTGGTGCACCGAGTGGTACATCGAACTGGATCTC
 AACAGCGTAAGATCCTGAGAGTTTCGCCCGAAGAACGTTTCAATGATGAGCACT
 TTTAAAGTCTGCTATGTGGCGGGTATTATCCCGTGTGACGCCGGCAAGAGCAACTC
 GGTGCCGCATAACACTATTCTCAGAATGACTTGGTGAGTACTCACCAAGTCACAGAAAAG
 CATCTAACGGATGGCATGACAGTAAGAGAATTATGCACTGCTGCATAACCATGAGTGA
 AACACTGCGGCCAACTTACTCTGACAACGATCGGAGGACCGAAGGGAGCTAACCGCTTT
 TTGCAACAACATGGGGATCATGTAACTCGCCTGATCGTTGGGAACGGAGCTGAATGAA
 GCCATACCAAACGACGAGCGTGACACCACGATGCCCTGAGCAATGGCAACAACGTTGCGC
 AAACATTAACGGCAACTTACTCTAGCTCCCGAACAAATTAAAGACTGGATG
 GAGGCGGATAAAAGTTGAGGACCACTCTCGCGCTCGGCTTCCGGCTGGTTATT
 GCTGATAAAATCTGGAGCGGTGAGCGTGGGTCTCGCGGTATATTGCACTGGGCA
 GATGGTAAGCCCTCCGTATCGTAGTTATCTACAGCAGGGAGTCAGGCAACTATGGAT
 GAACGAAATAGACAGATCGTGAAGATAGGTGCTCACTGATTAAGCATTGGTAACTGTCA
 GACCAAGTTACTCATATACTTTAGATTGATTAAACTCATTTAATTAAAGG
 ATCTAGGTGAAGATCCTTTGATAATCTCATGACCAAAATCCCTAACGTGAGTTTCC
 TTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTGAGATCCTTTT
 CTGCGCTAATCTGCTGCTGCAAACAAAAACCCACCGCTACAGCGGTGGTTGTTG
 CGGATCAAGAGCTACCAACTCTTCCGAAGGTAACGCTCAGCAGAGCGCAGATA
 CCAAATACTGCTCTAGTGTAGCCGTAGGCAACACTCAAGAAACTCTGTAGCA
 CGCCTACATACCTCGCTGCTAACCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAG
 TCGTGTCTTACCGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGAGCGTGGC
 TGAACGGGGGTTGTGACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGA
 TACCTACAGCGTGAAGCTATGAGAAAGGCCACGCTTCCGAAGGGAGAAAGGGGACAGG
 TATCCGTAAGCGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTCCAGGGGAAAC
 GCCTGGTATCTTATAGTCTGCTGGTTGCCCCCTCTGACTTGAGCGTCATTG
 TGATGCTCGTCAGGGGGCGGAGCCTATGGAAAACGCCAGCAACCGCCCTTTTACGG
 TTCCCTGCCCTTTGCTGCCCTTGCTCACATGTTCTTCTGCGTTATCCCTGATTCT
 GTGGATAACCGTATTACCGCCTTGAGTGAACCGCTCGCCAGCCGAACGACC
 GAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATCGGTATTCTCCTT
 ACGCATCTGCGGTATTCACACCGCATATATGGTCACTCTCAGTACAATCTGCTCTG
 ATGCCGCATAGTTAACCGACTATACACTCCGCTATCGCTACGTGACTGGTCATGGCTGC
 GCCCCGACACCCGCCAACACCGCTGACGCCCTGACGGGCTGCTGCCGGCATC
 CGCTTACAGACAACTGTGACCGTCTCCGGAGCTGCATGTGTCAGAGGTTTCAACCGTC
 ATCACCGAAACCGCGAGGAGCTGGTAAAGCTCATCGCTGGTGAAGCGATT
 ACAGATGCTGCCGTTACCGCGTCCAGCTCGTGAAGTTCTCCAGAACGCTTAATGT
 CTGGCTCTGATAAAAGCGGCCATGTTAAGGGCGTTTTCTGTTGGTACTGATGC
 CTCCGTTAAGGGGATTCTGTTATGGGTAATGATAACCGATGAAACGAGAGGAG
 GCTCACGATACGGTTACTGATGATGAAACATGCCGTTACTGGAACGTTGAGGGTAA
 ACAACTGGCGGTATGGATGCCGGGACAGAGAAAAACTACTCAGGGTCAATGCCAGCG
 CTTCGTTAACAGATGTTAGGTGTTCCACAGGGTAGCCAGCAGCATCTGCGATGCAGAT
 CGGAACATAATGGTGCAGGGCGCTGACTTCCCGTTCCAGACTTACGAAACACGGAA
 ACCGAAGACCATTGATGTTGCTCAGGTGCGAGACGTTTCCAGCAGCAGTCGCTTCA
 CGTTCGCTCGCGTATCGGTGATTCTGCTAACCGATGAGGCAAGGACCCGCCAGCCTAG
 CGGGGTCCCTCAACGACAGGAGCACGATCATGCCACCGTGGCCAGGACCAACGCTGCC
 CGAGATGCGCCGCGTGCAGGGCTGCTGGAGATGGCGGACGCGATGGATATGTTCTGCCAAGG
 GTTGGTTGCGCATTACAGTTCCGCAAGAATTGATTGGCTCCAATTCTGGAGTGGT-

FIGURE 96C

229/240

GAATCCGTTAGCGAGGTGCCGCCGGCTTCCATTCAAGTCAGGTGGCCCCGGCTCCATGCA
CCGCGACGCAACGCGGGGAGGCAGACAAGGTATAAGGGCGGCCCTACAATCCATGCCAAC
CCGTTCCATGTGCTCGCCGAGGCAGGCGCATAAATCGCCGTGACGATCAGCGGTCCAGTGATC
GAAGTTAGGCTGGTAAGAGCCGAGCGATCCTGAAGCTGTCCTGATGGTCGTATCT
ACCTGCCCTGGACAGCATGGCCTGCAACCGGGCATCCCGATGCCGAGCGAGAAGA
ATCATAATGGGAAGGCCATCCAGCCTCGCGTGCAGCACGCCAGCAAGACGTAGCCCAGC
GCGTCGGCCGCCATGCCGGCATAATGGCCTGCTCTGCCGAAACGTTGGTGGCGGGGA
CCAGTGACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCAATACCGCAAGCGACAGGCCG
ATCATCGTCGCGCTCCAGCGAAAGCGGTCCCTGCCGAAAATGACCCAGAGCGCTGCCGGC
ACCTGTCTACGAGTTGCATGATAAAAGAACAGTCATAAGTGCAGCGACGATAGTCATG
CCCCCGCCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGTCGATCG
ACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTT
GAGCACCGCCGCCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGC
CACGGGGCCTGCCACCATAACCCACGCCAAACAGCGCTATGAGCCGAAGTGGCGAGC
CCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCC
GGTGTGCCGGCACGATGCGTCCGGCGTAGAGG

FIGURE 96D

230/240

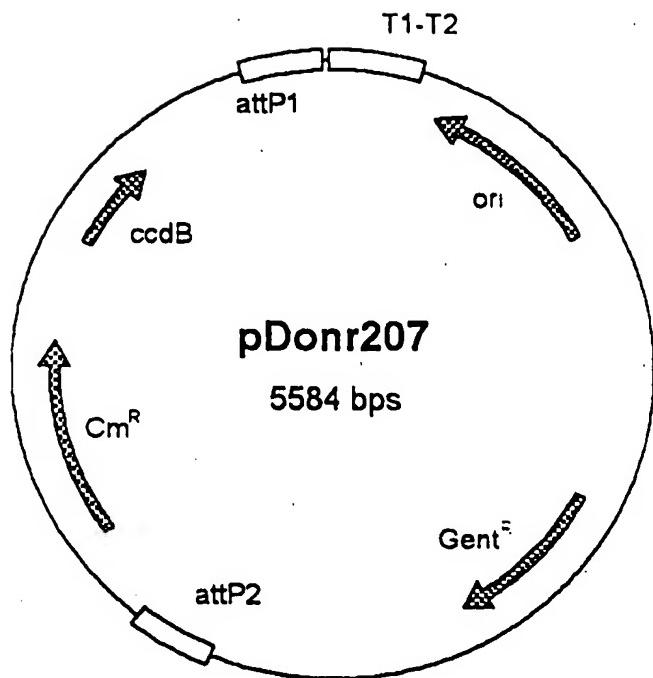


FIGURE 97A

pDONR207 5584 bp

GCGAGAGTAGGAACTGCCAGGCATCAAATAAAACGAAAGGCTAGTCGGAAAGACTGGC
 CTTTCGTTTATCTGTTGCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGG
 AGCGGATTGAACGTTGTGAAGCAACGCCGGAGGGTGGCGGGCAGGACGCCATA
 AACTGCCAGGCATCAAACTAAGCAGAAGGCCATCCTGACGGATGCCCTTTGCGTTCT
 ACAAAACTCTCCTGGCTAGCGGTAAACGGTTATCCACAGAATCAGGGATAACGCAGGA
 AAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAGGCCGCGTTGCTG
 GCGTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAG
 AGGTGGCGAAACCCGACAGGACTATAAGATAACCAGGCGTTCCCCCTGGAAGCTCCCTC
 GTGCGCTCTCCTGTTCCGACCCCTGCCGTTACCGGATACCTGTCGCCCTTCTCCCTCG
 GGAAGCGTGGCGTTCTCATAGCTACCGCTGTAGGTATCTCAGTTCGGTGTAGGTGTT
 CGCTCCAAGCTGGGCTGTGTCACGAACCCCCCTTCAGCCGACCCGCTGCCCTTATCC
 GGTAACTATCGTCTTGAGTCAACCCGTAAGACACGACTATCGCCACTGGCAGCAGCC
 ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTACAGAGTTCTGAAGTGG
 TGGCCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCCA
 GTTACCTCGGAAAAGAGTTGGTAGCTTGATCCGAAACAAACACCGCTGGTAGC
 GGTGGTTTTGTTGCAAGCAGCATTACGCGCAGAAAAAAAGGATCTCAAGAACGAT
 CCTTGATCTTCTACGGGTCTGACGCTCAGTGGAACGAAACACTCACGTTAAGGGATT
 TTGGTCACTGAGCTTGCCTCCCGTCAAGTCAGCGTAATGCTCTGCCAGTGTACAACC
 AATTAAACCAATTCTGATTAGAAAAACTCATCGAGCATCAATGAAACTGCAATTATTCA
 TATCAGGATTATCAATACCATATTGGAAAAGCCGTTCTGTAATGAAGGAGAAAAC
 CACCGAGGAGCTTCATAGGATGGCAAGATCCTGGTATCGTCTGCATTCCGACTCGTC
 CAACATCAATACAACCTATTAGTAGCCAACCACTAGAACTATAGCTAGAGTCCTGGCGA
 AAAAAACGATGCTGCCCTCCAGAAAACCGAGGATGCGAACCAACTTCATCCGGGTAGCA
 CCACCGGCAAGGCCCGCAGGCCGAGGTCTCCGATCTCTGAAAGCCAGGGCAGATCCG
 TGACAGCACCTGCCGTAGAAGAACAGCAAGGCCCAATGCCCTGACGATGCCGTGGAGA
 CCGAAACCTTGCGCTGCCAGGCCAGGACAGAAATGCTCGACTTCGCTGCTGCCCA
 AGGTTGCCGGTGACGACACCGTGGAAACGGATGAAGGCACGAACCCAGTTGACATAAG
 CCTGTTCCGTTCTGAAACTGTAATGCAAGTAGCGTATGCCCTCACGCAACTGGTCCAGAA
 CCTTGACCGAACGCCAGCGGGTAAAGGCGCAGTGGCGTTTCATGGCTGTTATGACT
 GTTTTTTGTCAGTCTATGCCCTGGCATCCAAGCAGCAAGCGCTTACGCCGTGGTC
 GATGTTGATGTTATGGAGCAGCAACCGATGTTACGCGACGCAACGATGTTACCGCAG
 GGCAGTCGCCCTAAAACAAAGTTAGGGCTCAAGTATGGCATTCGCACATGTAAG
 CTCCGCCCCCTGACCAAGTCAAATCCATGCCGCTGCTCTGATCTTCGGTCTGAGTT
 GGAGACGTAGCCACCTACTCCAAACATCAGCCGACTCCGATTACCTGGGAACCTGCTC
 CGTAGTAAGACATTCATGCCCTGCTGCCCTGCCAGAAGACGGTTGTTGGCGCTCTC
 GCGCTTACGTTGCCAGGTTGAGCAGCCGCTAGTGAGATCTATATCTATGATCTC
 GCAGTCTCCGGCGAGCACGGAGGCAGGGCATTGCCACCAGCCTCATCAATCTCCTCAAG
 CATGAGGCCAACGCCCTGGTCTATGTGATCTACGTGCAAGCAGATTACGGTACGAT
 CCCGAGTGGCTCTATACAAAGTTGGCATAACGGGAAAGAAGTGTACGACTTTGATATC
 GACCCAAAGTACCGCCACCTAACATTGTTCAAGCCGAGATCGGCTTCCGGCTAATTT
 CCCCTCGTCAAAATAAGGTTATCAAGTGAGAAATCACCAGAGTGACGACTGAATCCGG
 TGAGAATGGAAAAGTTATGCAATTCTTCCAGACTTGTCAACAGGCCAGCATTACG
 CTCGTCATCAAAATCACTCGCATCAACCAACCGTTATTCACTGATGCGCCTGAGC
 GAGACGAAATACCGGATCGCTTTAAAGGACAATTACAAACAGGAATGCAATGCAACCG
 GCGCAGGAACACTGCCAGCCATCAACAAATATTTCACCTGAAATCAGGATATTCTCTAA
 TACCTGGAATGCTGTTTCCGGGATCGCAGTGGTGTAGTAACCATGCGATCATCAGGAGT
 ACGGATAAAATGCTGATGGCGGAAGAGGCATAAAATTCCGTACGCCAGTTAGTCTGAC
 CATCTCATCTGTAACATCATTGGCAACGCTACCTTGCCATGTTCAAGAAACAACTCTGG
 CGCAGTCCGGCTTCCATACAAGCGATAGATTGTCGCACCTGATTGCCGACATTATCGCG
 AGCCCATTATACCCATATAAAATCAGCATCCATGTTGAATTAAATCGCGGCCCTGACGT
 TTCCCGTTGAATATGGCTCATACACACCCCTGTATTACTGTTTATGTAAGCAGACAGTT
 TATTGTTCACTGATGATATTTTATCTTGTCGAACTGAAACATCAGAGATTGGAGACAC
 GGGCCAGAGCTGCCAGCTGGATGGCAAAATAATGATTTTATTGACTGATAGTGACCTGTT
 CGTTGCAACAAATTGATAAGCAATGCTTCTATAATGCCAACTTTGTACAAGAAAGCTG
 AACGAGAAACGTAATGATATAATCAATATATTAAATTAGATTTGCATAAAAAC
 AGACTACATAACTGTAACACACACATCCAGTCACATGAACTACTTAGATG-

FIGURE 975

GTATTAGT GACCT GTAGTCGACTAAGITGGCAGCATCACCCGACGCAC T T GCGCCGAAT
AAATAACCTGTGACGGAAGATCACTTCG CAGAATAAATAAATCTGGTGTCCCTGTTGATA
CCGGGAAGCCCTGGCCAAC T TGGC GAAAATGAGACGTTGATCGGCACGTAAGAGGTTC
CAACTTTCAACCATAATGAAATAAGATCACTACCGGGCGTATTTTGAGTTATCGAGATT
TTCAGGGAGCTAAGGAAGCTAAATGGAGAAAAAAATCACTGGATATACCACCGTTGATAT
ATCCC AATGGCATCGTAAAGAACATTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTA
TAACCCAGACC GTT CAGCTGGATATTACGCCCTTTAAAGACCGTAAGAAAAATAAGCA
CAAGTTTATCCGGCCTTATTACATTCTGCCCGCTGATGAATGCTCATCCGAATT
CCGTATGGCAATGAAAGACGGTGAGCTGGT GATATGGGATAGTGTTCACCCCTGTTACAC
CGTTTCCATGAGCAA ACTGAAACGTTTCATCGCTCTGGAGTGAAATACCACGACGATT
CCGGCAGTTCTACACATATATTGCAAGATGTGGCGTGTACGGT GAAAACCTGGCCTA
TTCCCTAAAGGGTTATTGAGAATATGTTTCTGCTCAGCCAATCCCTGGGTGAGTT
CACCAGTTT GATTTAAACGTGGCAATATGGACAAC T T CTCGCCCGTTTCACCAT
GGCAAAATATTACGCAAGCGACAAGGTGCTGATGCCGCTGGCATT CAGGTTCATCA
TGCCGTCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGA
TGAGTGGCAGGGCGGGCGTAA TCGCGTGGATCCGGCTACTAAAAGCCAGATAACAGTA
TGC GTATTGCGCGCTGATT T T GCGGTATAAGAATATACTGATATGTATACCCGAAG
TATGTCAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTGACAGCGACAGC
TATCAGTTGCTCAAGGCATATATGATGTC AATATCTCCGGTCTGGTAAGCACAACCATGC
AGAATGAAGCCC GTCGTCTGCGTGCCAAGCTGGAAAGCGGAAAATCAGGAAGGGATGG
CTGAGGTGCGCCGGTTATTGAAATGAACGGCTTTGCTGACGAGAACAGGACTGGT
GAAATGCAGTTAAGGTTACACCTATAAAGAGAGAGCCGTTATCGTCTGTTGTGGAT
GTACAGAGTGATATTATTGACACGCCCGGGCAGGGATGGT GATCCCCCTGGCAGTGCA
CGTCTGCTGTCAAGATAAGTCTCCCGTGAAC T TACCCGGTGGTGCATATGGGGATGAA
AGCTGGCGCATGATGACCACCGATATGCCAGTGTGCCGGTCTCGTTATGGGAAGAA
GTGGCTGATCTAGCCACCGC GAAAATGACATCAAAACGCCATTACCTGATGTTCTGG
GGAATATAAATGTCAGGCTCCCTTATACACAGCCAGTCTGCAAGTGATACAGTAGAAAT
TACAGAAACTTATCAGTTAGTAAGTATAGAGGCTGAAAATCCAGATGAAGCCGAACG
ACTTGTAAAGAGAAAAGTATAAGAGTTGTAAGATTGTTCTGATGCA GATGATTTCAGGA
CTATGACACTAGCGTATATGAATAGGTAGATGTTTATTGTCACACAAAAAAGAGGC
TCGCACCTCTTTCTTATTCTTATGATTAA TACGGCATTGAGGACAATAGCGAG
TAGGCTGGATACGACGATTCCGTTGAGAAGAACATTGGAAGGCTGCGGTGACTAAG
TTGGCAGCATCACCAGAACATTTGGAAGGCTGCGGTGACTACAGGTCACTAATAC
CATCTAAGTAGTTGATTCAAGTGACTGGATATGTTGTTACAGTATTATGTAGTCT
GTTTTTATGCAAATCTAATTAA TATATTGATATTATCATTACGTTCTCGTT
CAGCTTTTGTA CAAAGTGGCATTATAAAAAGCATTGCTCATCAATTGTTGCAACG
AACAGGTCACTATCAGTCAA AATAATCATTATTGGGCCCGAGATCCATGCTAGCGT
TAAC

FIGURE 97C

233/240

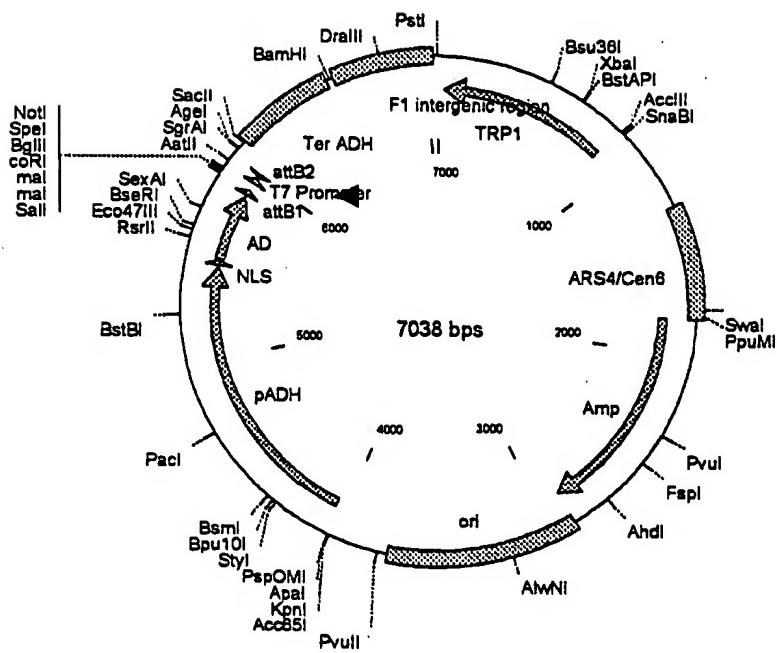
pMAB85

FIGURE 98A

pMAB85 7038 bp

GCCTTACGCATCTGTGCGGTATTCACACCGCAGGCAAGTCACAAACAATACTTAAATA
 AATACTACTCAGTAATAACCTATTCTTAGCATTTGACGAAATTGCTATTGTTAG
 AGTCTTTACACCATTGCTCCACACCTCCGCTTACATCAACACCAATAACGCCATT
 ATCTAAGCGCATACCAACATTCTGGCGTCAGTCCACCAAGCTAACATAAAATGTAAGC
 TTTCGGGCTCTCTGGCTTCAACCCAGTCAGAAATCGAGTTCCAATCCAAAAGTCAC
 CTGTCCCACCTGCTCTGAACAAACAGGAATAAACGAATGAGGTTCTGTGAAGCTG
 CACTGAGTAGTATGCGAGCTTTGGAAATACGAGTCCTTAATAACTGGCAAACCGA
 GGAACCTTGGTATTCTTGCCACGACTCATCTCCATGCAGTGGACGATATCAATGCCG
 AATCATTGGCAGAGCCAAACATCCCTTAGGTTGATTACGAAACACGCCAACCAAGT
 ATTCGGAGTGGCTGAACATTTTATATGCTTTACAAGACTGAAATTTCCTGCAA
 TAACCGGGTCAATTGTCCTTCTATTGGGCACACATATAATACCCAGCAAGTCAGCAT
 CGGAATCTAGAGCACATTCTGGCCCTCTGTGCTCTGCAAGCCGAAACTTCACCAATG
 GACCAGAACTACCTGTGAAATTAAATAACAGACATACTCCAAGCTGCTTGTGCTTAA
 TCACGTATACTCACGTGCTCAATAGTACCAATGCCCTCCCTTGGCCCTCTCCTTT
 TTTTCGACCGAATTAATTCTTAATCGGAAAAAGAAAAGCTCCGGATCAAGATTG
 ACCTGAAAGTACAAGCTATTTCATAAAAGAATATCTTCAACTACTGCCATCTGGCGTC
 ATAACGTCAAAGTACACATATAATTACGATGCTGTCTATTAAATGCTCCTATATTATA
 TATAGTAATGCTTTATGGTGCACCTCAGTACAATCTGCTCTGATGCCGATAGTTAA
 GCCAGCCCCGACACCCGCAACACCCGCTGACCGCCCTGACGGGCTTGTCTGCTCCC
 CATCCGCTTACAGACAAGCTGTGACCGTCTCCGGAGCTGCATGTGTCAGAGGTTTC
 CGTCATCACGAAACCGCGAGACGAAAGGGCCTCGTGTACGCCTATTAGTTAGGTTA
 ATGTCATGATAATAATGGTTCTTAGGACGGATCGCTGCTGTAACTTACACGCCCTC
 GTATCTTTAATGATGGAATAATTGGGAAATTACTCTGTGTTATTATTAGTTAGTT
 TGTATTGGATTAGAAAGTAAATAAGAAGTAGAAGAGTTACGGAATGAAGAAAAAA
 AAATAACAAAGGTTAAAAAATTCAACAAAAGCTACTTACATATAATTAGTTAG
 ACAAGAAAAGCAGATTAAATAGATATACATTGATTAACGATAAGTAAATGAAAATCA
 CAGGATTTCTGTGTGGCTTACACAGACAAGATGAAACAATTGGCATTAATACCT
 GAGAGCAGGAAGAGCAAGATAAAAGTAGTATTGTTGGCATTCCCTAGAGTCTTTA
 CATCTCGAAAACAAAACATTCTTCTTAAATTCTTTTACTTTCTATTAA
 TTTATATATTATATAAAATTAAATTATAATTAGCACGTGATGAAAAG
 GACCCAGGTGGCATTTCGGGAAATGTGCGCGAACCCCTATTGTTATTCTAA
 ATACATTCAAATATGTATCCGCTCATGAGACAATAACCCGATAAATGCTTCAATAAT
 TGAAAAGGAAGAGTATGAGTATTCAACATTCCGTGTCGCCCTTATTCCCTTTGCG
 GCATTTCGCTTCTGTTGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAA
 GATCAGTTGGGTGACGAGTGGTTACATCGAATGGATCTCAACAGCGGTAGATCCT
 GAGAGTTTCGCCCGAAGAACGTTCCAAATGATGAGCAGCTTTAAAGTTCTGCTATGT
 GGCGCGTATTATCCGTATTGACGCCGGCAAGAGCAACTCGGTGCGCATACACTAT
 TCTCAGAATGACTTGGTTGAGTACTCACCAAGTCACAGAAAAGCATTTACGGATGGC
 ACAGTAAGAGAAATTGCACTGCTGCCATAACCATGAGTATAACACTGGGCCAACTTA
 CTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTCAACACATGGGG
 CATGTAACCTGCCCTTGATCGTTGGGAAACCGGAGCTGAATGAAGCCATACCAACGAC
 CGTACACCAACGATGCCGTAGCAATGGCAACACGTTGCGCAAACATATTAAACTGG
 CTACTACTCTAGCTCCCGAACAAATAATAGACTGGATGGAGGCGGATAAGTTGCA
 GGACCACTCTCGCCTCGGCCCTCCGGCTGGCTGGTTATTGCTGATAAAATCTGGAG
 GGTGAGCGTGGGTCTCGGGTATCTGAGCAGCACTGGGCCAGATGTTAGCCCTCC
 ATCGTAGTTATCACGACGGGAGTCAGGAACTATGGATGAACGAAATAGACAGATC
 GCTGAGATAGGTGCTCACGATTAAGCATTGGTAACTGTCAGACAAGTTACTCAT
 ATACTTTAGATTGATTAAACTTCATTAAATTAAAGGATCTAGGTGAAGATCCTT
 TTTGATAATCTCATGACCAAAATCCCTAACGTGAGTTTCGTTCACTGAGCGTCAGAC
 CCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTCTGCGCGTAATCTGCTG
 TTGCAAACAAAAACACCGTACCGAGCGGTGGTTGTTGCGGATCAAGAGCTACCA
 ACTCTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATAACAAATACTGCTCTA
 GTGTAGCCGTAGTTAGGCCACCACTCAAGAATCTGAGCACCCTACATACCTCGCT
 CTGCTAACCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGG
 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGCTGAACGGGGTTCGTGC

FIGURE 9B

235 / 240

ACACAGCCCAGCTGGAGCGAACGACCTACACCGAACTGAGATACTACAGCGTGAGCAT
 TGAGAAAAGGCCACGCTTCCGAAGGGAGAAAGCGGACAGGTATCCGTAAGCGGCAGG
 GTCGGAACAGGAGAGCGCACGAGGGAGCTCCAGGGGGAACGCCCTGGTATCTTATAGT
 CCTGTGGGTTTCGCCACCTCTGACTTGAGCGTCATTGTGATGCTCGTCAGGGGG
 CCGAGCCTATGGAAAAACGCCAGCAACCGGCCCTTTACCGTTCTGGCCTTTGCTGG
 CCTTTGCTCACATGTTCTTCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACC
 GCCTTGAGTGGAGCTGATACCGCTGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTG
 AGCGAGGAAGCGGAAGAGCGCCAAACGCAAACCGCCTCTCCCGCGCGTTGCCGATT
 CATTAATGCAGCTGGCACGACAGGTTCCGACTGGAAAGCGGGAGTGGAGCGCAACGCA
 ATTAATGTGAGTTACCTCACTCATTAGGCACCCAGGCTTACACTTATGCTCCGGCT
 CCTATGTTGTTGGAATTGTGAGCGGATAACAATTACACAGGAACAGCTATGACCAT
 GATTACGCCAAGCTCGGAATTAAACCTCACTAAAGGAACAAAGCTGGGTACCGGGCC
 CCCCTCGAGATCCGGATCGAAGAAATGATGGTAAATGAAATAGGAATCAAGGAGCATG
 AAGGCAAAAGACAAATATAAGGGTCGAACGAAAAATAAGTGAAGAAGTGTGATATGATG
 TATTGCTTGCAGCCGCCCCGGCGATAACGCTGGCGTGGAGCTGTGCCCCGGC
 GTGGCGACCCCGCTTGCAGCCGCGATAACGCTGGCGTGGAGCTGTGCCCCGGC
 GGAGTTTGCAGCCGCTGATTTCCAAGGTTACCGCTGCGCTAAGGGGGAGATTGGAGA
 AGCAATAAGAATGCCGGTTGGGTTGCGATGATGACGACACGACAACACTGGTGTCTTAT
 TTAAGTGGCGAAAGAACCTGAGTGCATTGCAACATGAGTATACTAGAAGAACGCA
 AGACTTGCAGACCGAGTTGCCGGTGGCGAACATAGAGCGGACATGACCTTGAAG
 GTGAGACGCGCATAACCGCTAGAGTACTTGAAGAGGAACAGCAATAGGTTGCTACCA
 GTATAAATAGACAGGTACATACAACACTGGAAATGGTTGCTGTTGAGTACGCTTCAA
 TTCATTGGGTGCACTTATTATGTTACAATATGGAAGGGAACTTACACTTCTCCTA
 TGCACATATATTAATTAAAGTCAAATGCTAGTAGAGAACAGGGGTAACACCCCTCCGCG
 TCTTTCCGATTTCTAAACCGTGGAAATTTCGGATATCCTTGTGTTCCGG
 TGTACAATATGGACTCCTCTTCTGGCAACCAACCCATACTCGGATTCTTCTATAAT
 ACCTTCGTTGGCTCCCTAACATGTAGGTGGCGAGGGAGATATAACATAGAACAGATA
 CCAGACAAGACATAATGGCTAAACAAAGACTACACCAATTACACTGCCTCATGATGGTG
 GTACATAACGAATAACTGTAGCCCTAGACTGTAGACCGCATCATATCGAAGTTTC
 ACTACCCCTTCCATTGCCATCTATTGAAGTAATAATAGGCGCATGCAACTTCTTTC
 TTTTTTTCTTCTCTCTCCCCGTTGTTGCTCACCATATCCGAATGACAAAAAAA
 ATGATGGAAGACACTAAAGAAAAATAACGACAAAGACAGCACCAACAGATGTCGTTG
 TTCCAGAGCTGATGAGGGGTATCTCGAACACACGAAACTTTCTCTCCTTCTTCA
 CACACTCTCTAACATGAGCAACGGTATACGGCCTCCTCAGTTACTGAATTGAAA
 TAAAAAAAGTTGCCGCTTGCTATCAAGTATAATAGACCTGCAATTATAATCTTTG
 TTTCCTCGTCATTGTTCTCGTCCCTTCTCTGTTCTTCTGACAATATTCA
 AGCTATACCAAGCATACAATCAACTCCAAGCTATGCCAACAGAACAGGGAGGTCTCG
 AGCGCGCAATTAAATCAAAGTGGAAATTGCTGATAGCTATTGCTCTCACTTC
 ACTAACAGTAGCAACGGTCCGAAACCTCATACAAACTCAAACAAATTCTCAAGCGCTTCA
 CAACCAATTGCCCTCTAACGTTCATGATAACTCATGAAATAATGAAATCACGGCTAGT
 AAAATTGATGATGGAATAATTCAAACCAACTGTGACCTGGTGGACGGACCAAACGCG
 TATAACCGTTGGAAATCACTACAGGGATGTTAACCAACTACAATGGATGATGTATAT
 AACTATCTATTGATGATGAAAGATAACCCACCAACAAAAAGAGGGTGGTCGATC
 ACAAGTTGTACAAAAAGCAGGCTGTCGACCCGGAAATTGAGACTACTAGTGCAGC
 CGCACCGTACCCAGCTTCTGACAAAGTGGTGACGTCAGCTCCCTATAGTGAAGTC
 TATTACACTGGCCGTCGTTTACAACGTCGTGACTGGAAAACACCGTGAGCTCAAGT
 AAGTAACGGCCGCCACCGCGGTGGAGCTTGGACTTCTCGCCAGAGGTTGGTCAGTC
 TCCAATCAAGGTTGTCGGCTTGTCTACCTTGCAGAAATTACGAAAAGATGGAAAAGGG
 TCAAATCGTTGGTAGATACTGTTGACACTTCTAAATAAGCAATTCTTATGATTAT
 GATTTTATTATAAAGTTATAAAAAAAATAAGTGTATACAATTAAAGTGAAC
 TTAGGTTTAAACGAAAATTCTTGGTCTTGAGTAACCTTCTGAGGTCAAGGTTGCT
 TTCTCAGGTATAGCATGAGGTGCGCTTATTGACCACACCTCTACCGCATGCCAGCA
 ATGCCTGCAAATCGCTCCCATTTACCCAAATTGAGATATGCTAATCCAGCAATGAGT
 TGATGAATCTCGGTGTGTTATGCTCTCAGAGGACAATACCTGTTGTAATCGTTCTT
 CCACACGGATCCGCATCAGCGAAATTGAAACGTTAAATTTGTTAAAATTGCGTTA
 AATATTGTTAAATCAGCTCATTTTAAACCAATAGGCGAAATCGGCAAAATCCCTTAT
 AAATCAAAGAACGAGATAGACCGAGATAGGGTTGAGTGTGTCAGTTGGAACAAGAGTCCA
 CTATTAAGAACGTGGACTCCAACGTCAAAGGGCGAAAACCGTCTATCAGGGCGATGGC-

FIGURE 9BC

236 / 240

CCACTACGTGAACCCTAACCCCTAATCAAGTTTTGGGTCGAGGTGCCGTAAAGCACTA
AATCGGAACCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGAAAGCCGGCGAACGTG
GCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGTAGGGCGCTGGCAAGTGTAGCG
GTCACGCTGCCTAACCAACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCC
CATTGCCATTCACTGCA

FIGURE 98D

237/240

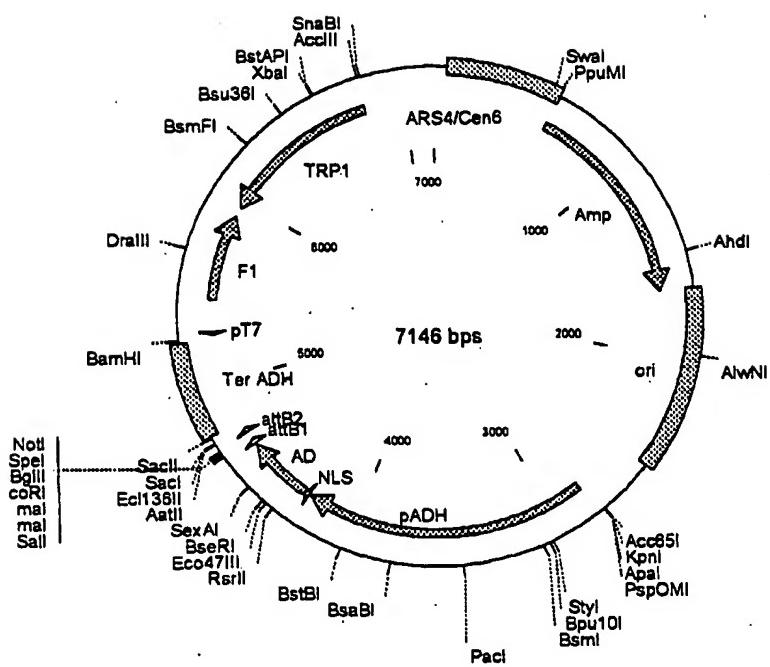
pMAB86

FIGURE 99A

pMAB86 7146 bp

GACGAAAGGGCCTCGTATACGCCTATTAGTTAATGTCATGATAATAATGGTT
 CTTAGGACGGATCGCTTGCCTGTAACCTACAGCGCCTCGTATCTTTAATGATGGAATA
 ATTTGGGAATTACTCTGTGTTATTATTTATGTTGATTTGGATTTAGAAAGT
 AAATAAAGAAGGTAGAAGAGTTACGGAATGAAGAAAAAAAATAACAAAGGTTAAAAA
 ATTTCAACAAAAGCGTACTTACATATATTTATTAGACAAGAAAAGCAGATTAATA
 GATATACATTGATTAACGATAAGTAAATGTAACAGGATTTCTGTGTGGTCT
 TCTACACAGACAAGATGAAACAATTCCGGCATTAACCTGAGAGCAGGAAGAGCAAGATA
 AAAGGTAGTATTGTTGGCGATCCCCTAGACTTTACATCTCGGAAACAAACAAACT
 ATTTTTCTTAATTCTTTACTTCTATTAAATTATTTATTTATTTAAAAA
 ATTTAAATTATAATTATTTATAGCACGTGATGAAAAGGACCCAGGTGGCACTTTGG
 GGAAATGTGCGCGAACCCCTATTGTTATTCTAAATACATTCAAATATGTATCCG
 CTCATGAGACAATAACCCTGATAAAATGCTTCAATAATATTGAAAAGGAAGAGTATGAGT
 ATTCAACATTCCGTGTCGCCCTTATTCCCTTTTGCGGCATTTGCCCTGTTT
 GCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAGATCAGTGGGTGACGAGTG
 GGTTACATCGAACTGGATCTAACAGCGGTAAAGATCCTTGAGAGTTTCGCCCGAAGAA
 CGTTTCCAATGATGAGCACTTTAAAGTTCTGCTATGTGGCGCGTATTATCCCGTATT
 GACGCCGGCAAGAGCAACTCGGTGCCGCATAACTATTCTCAGAAATGACTTGGTGAG
 TACTCACCAAGTCACAGAAAAGCATCTAACGGATGGCATGACAGTAAGAGAATTATGCAGT
 GCTGCCATAACCCTGAGTGATAACACTCGGCCAACTTACTCTGACAACGATCGGAGGA
 CGGAAGGAGCTAACCGCTTTTCAACATGGGGATCATGTAACTCGCCTGATCGT
 TGGGAACCGGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGCACACCACGATGCCGT
 GCAATGGCAACACCGTGGCAAACACTTAACCTGGCAACTACTACTCTAGCTTCCCG
 CAACAATTAAAGACTGGATGGAGGGGATAAAGTTGAGGACCACTCTGCCTCGGCC
 CTTCGGCTGGCTGGTTATTGCTGATAAAATCTGGAGCCGTGAGCGTGGGTCTCGCGGT
 ATCATTGCAGCACTGGGCCAGATGGTAAGCCCTCCGTATCGTAGTTATCTACACGACG
 GGCAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCTCACTG
 ATTAAGCATTGGTAACGTGACGACAAAGTTACTCATATAACTTTAGATTGATTAAAAA
 CTTCATTAAATTAAAGGATCTAGGTGAAGATCTTGTATAATCTCATGACCAAA
 ATCCCTTAACGTGAGTTTCGTTCACTGAGCGTCAGACCCGTAGAAAAGATCAAAGGA
 TCTTCTTGAGATCTTTCTGCGCGTAACTGCTGCTGAAACAAAAACCAACCG
 CTACCAAGCGGTGGTTGCGGATCAAGAGCTACCAACTCTTCCGAAGGTAAC
 GGCTTCAGCAGAGCGCAGATACCAAAACTGTCTCTAGTGTAGCCGTAGTTAGGCCAC
 CACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTGCTAATCCGTACCGAGTG
 GCTGCTGCCAGTGGCGATAAGTCGTCTTACGGGTTGGACTCAAGACGATAGTTACCG
 GATAAGGCGCAGCGGTGGCTGAACGGGGGTTCTGTGCACACAGCCCAGCTGGAGCGA
 ACGACCTACACCGAACTGAGATACTACAGCGTGAGCATTGAGAAAGCGCCACGCC
 GAAGGGAGAAAGGCCAGGTATCCGTAAAGCGGCAGGGTGGAAACAGGAGAGCGCAGG
 AGGGAGCTCCAGGGGGAAACGCGTGGTATCTTATAGTCTGCGGTTTCGCCACCTC
 TGACTTGAGCGTCGATTTGTGATGCTCGTCAGGGGGCCGAGCCTATGAAAAACGCC
 AGCAACCGGGCTTTACGGTCTGGCTTTGCTGGCTTTGCTCACATGTTCTT
 CCTGCGTTATCCCTGATTGTTGATAACCGTATTACCGCCTTGAGTGAGCTGATACC
 GCTCGCCGAGCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGC
 CCAATACGCAAACCGCTCTCCCCCGCTGGCGATTCTTAATGCACTGGCACGAC
 AGGTTCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTATGTGAGTTACCTCACT
 CATTAGGCACCCAGGTTACACTTATGCTTCCGGCTCTATGTTGTTGGAATTG
 AGCGGATAACAATTACACAGGAAACAGCTATGACCATGATTACGCCAAGCTCGGAATT
 AACCCCTACTAAAGGAAACAAAGCTGGTACCGGGCCCCCTCGAGATCCGGATCGA
 AGAAATGATGGTAATGAAATAGGAAATCAAGGAGCATGAAGGCAAAGACAAATATAAG
 GTGCGAACGAAAATAAAAGTGAAGAGTGTGATATGATGTTGAGCTGAGGCGCCGA
 AAAACGAGTTACGCAATTGACAAATCATGCTGACTCTGTGGCGACCCCGCTTGC
 CGGCCCCGGCGATAACGCTGGCGTGAGGCTGTGCCCGGGAGTTTGTGCGCTGCATT
 TTCCAAGGTTACCTGCGCTAAGGGCGAGATTGAGAAGCAATAAGAATGCCGGTTGG
 GGTTGCGATGATGACGACCAAGACAACCTGGTGTCTTAAAGTTGCGAAAGAACCTG
 AGTGCATTGCAACATGAGTAACTAGAAGAATGAGCAAGACTTGCGAGACGCGAGTT
 GCCGGTGGTGCAGAACATAGAGCGACCATGACCTTGAAGGTGAGACGCGCATAACCGCTA-

FIGURE 99B

239/240

GAGTACTTGAAAGAGGAAACAGCAATAGGGTGTACCGATATAAATAGACAGGTACATA
 CAACACTGGAAATGGTGTCTGTTGAGTACGCTTCATTGAGGTGTGACTTTA
 TTATGTTACAATATGGAAGGGAACCTTACACTCTCCTATGCACATATAATTAAAGT
 CCAATGCTAGTAGAGAAGGGGGTAACACCCCTCCGCGCTTTCCGATTTCTAA
 ACCGTGGAATATTCGGATATCCTTGTGTTCCGGGTACAATATGGACTTCCTCT
 TTTCTGGCAACCAAACCCATACATCGGGATTCCATAATAACCTTCGTTGGTCTCCCTAAC
 ATGTAGGGGGAGGAGATAACATAGAACAGATACCAGAACAGACATAATGGGCT
 AAACAAGACTACACCAATTACACTGCCTCATGATGGTGTACATAACGAACAACTAATCTG
 TAGCCCTAGACTGATAGCCATCATCATCGAAGTTCACTACCCCTTTCCATTGCC
 ATCTATTGAAGTAATAATAGCGCATGCAACTCTTTCTTTCTCTCTC
 CCCCGTTGTTGTCACCATATCCGCAATGACAAAAAAATGATGGAAGACACTAAAGGA
 AAAAATTAAACGACAAAGACAGCACCAACAGATGTCGTTCCAGAGCTGATGAGGGTA
 TCTTCGAACACACGAAACTTTCTCTCATTACGCACACTACTCTATAATGAGCA
 ACGGTATACGGCCCTCCTCCAGTTACTGAAATTGAAATAAAAAAGTTGCCGTTG
 CTATCAAGTATAAAATAGACCTGCAATTATAATCTTTGTTCTCGTCTCGT
 TCCCTTCTCCTGTTCTTCTGCACAATATTCAGCTATAACGACATAACATC
 AACCTCCAAGCTTATGCCAAGAAGAGCGGAAGGTCTCGAGCGGCCAATTAAATCAA
 AGTGGGAATATTGCTGATAGCTCATTGCTCTCACTTCAACAGTAGCAACGGTCCG
 AACCTCATACAACTCAAACAAATTCTCAAGCGCTTCAACACATTGCCCTCTAAC
 GTTCATGATAACTCATGAAATAATGAAATCACGGTAGTAAATTGATGATGGTAAAT
 TCAAAACCACTGTCACCTGGTGGACGACAAACTGCGTATAACCGTTGGAATCACT
 ACAGGGATGTTAATACCAACTACAATGGATGATGATAATAACTATCTATTGATGATGAA
 GATACCCCACCAACCAAAAAAGAGGGTGGGTGATCAGATTGTTGACAAAAGCA
 GGCTTGTCGACCCGGGAATTAGCTACTAGTGCAGCCGACCGCTACCCAGCTTCT
 TGTACAAAGTGGTGCAGTCGAGCTCTAGTAAGTAACGGCCGCCACCGGGTGGAGCTT
 GGACTTCTCGCCAGAGGTTGGTCAAGTCTCAATCAAGGTTGTCGGCTGCTACCTT
 GCCAGAAATTACGAAAAGATGAAAGGGTCAATCGTTGGTAGATACGTTGTTGACAC
 TTCTAAATAAGCAATTCTTATGATTATGATTTTATTAAATAAGTTATAAAAAAA
 AATAAGTGTATAACAAATTAAAGTACTCTAGGTTAAAACGAAAATTCTGTTCTT
 GAGTAACCTTCTGTAGGTCAGGTTGCTTCTCAGGTATAGCATGAGGTGCTCTTAT
 TGACCCACACTCTACCGGCATGCCGAGCAAATGCCGCAAAATCGCTCCCCATTCAACCA
 ATTGTAGATATGCTAACTCCAGCAATGAGTTGATGAATCTCGGTGTTATTTATGCT
 CAGAGGACAATACTGTTGTAATCGTCTTCCACACGGATCCAACTGCCCTATAGTGA
 GTCGTATTACAATTCACTGCCGTCGTTTACAACGTCGACTGGAAAACCTGGCGT
 TACCCAACTTAATCGCCTTGCAAGCACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGA
 GGCCCGACCGATGCCCTCCAAACAGTTGCCAGCCTGAATGGCAATGGACGCC
 TGTAGCGCGCATTAAAGCGGGGGGTGTTACGCCAGCGTACCGCTACACT
 GCCAGGCCCTAGGCCGCTCTTCGCTTCTCCCTCTTCGCCACGTTGCC
 GGCTTCCCCGTCAGCTCAAATCGGGGCTCCCTTAGGGTCCGATTAGTGT
 CGGCACCTCGACCCAAAAAACTGATTAGGGTGTACGTTAGGGCCATGCC
 TGATAGACGGTTTCGCCCTTGACGTTGGAGTCCACGTTCTTAAAGTGGACTCTG
 TTCCAAACTGGAACACACTCAACCCATCTGGTCTATTCTTTGATTATAAGGGATT
 TTGCCGATTCGCCATTGGTAAAAAATGAGCTGATTAAACAAAATTAAACGCGAAT
 TTTAACAAAATTAAACGTTACAATTCTGATGCCGTTATTCCTTACGCATCTGT
 CGGGTATTCACACCGCAGGCAAGTCACAAACAAATACTAAATAACTCAGTAA
 TAACCTATTCTTAGCATTGGACGAAATTGCTATTGTTAGAGTCTTACACC
 TTGTCTCCACACCTCGCTACATCAACACCAATAACGCATTAACTAAGCGCATT
 CAACATTCTGGCGTCAGCCACCGCTAACATAAAATGTAAGCTTCCGGCTCTT
 GCCTTCCAACCCAGTCAGAAATCGAGTTCCAATCCAAAAGTTCACCTGCCC
 CTGAAATCAAACAGGAATAACGAATGAGGTTCTGTGAGCTGCACTGAGTAGTATGT
 TGCAGTCTTGGAAATACGAGTCTTTAATAACTGGCAACCGAGGAACCTGGTATT
 CTTGCCACGACTCATCTCATGCACTGGACGATATCAATGCCGTAATCATTGACCAGAG
 CCAAAACATCCTCCTTAGGGTGTACGAAACAGGCCACCAAGTATTGCGAGTGCCTG
 AACTATTATGCTTTACAAGACTGAAATTCTTGCATAACCGGGTCAATTG
 TTCTCTTCTATTGGGACACATATAACCCAGCAAGTCAGCATCGGAATCTAGAGCAC
 ATTCTGCCGCTCTGTGCTGCAAGGCCAACTTTACCAATGGACCAGAACACTACCG
 TGAAATTAAACAGACATACTCCAAGCTGCCCTTGCTGTTAATCACGTAACTCAG
 TGCTCAATAGTCACCAATGCCCTCCCTTGCCCCCTCTCTTTTCGACCGAAT-

FIGURE 9c

240 / 240

TAATTCTTAATCGGCAAAAAAGAAAAGCTCCGGATCAAGATTGTACGTAAGGTGACAAG
CTATTTTCATAAAGAATATCTTCCACTACTGCCATCTGGCGTCATAACTGCAGTAC
ACATATATTACGATGCTGTCTATTAAATGCTTCTATATTATATATAGTAATGTCGTT
TATGGTGCACCTCAGTACAATCTGCTCTGATGCCGATAGTTAAGCCAGCCCCGACACC
GCCAACACCCGCTGACGCCCTGACGGGCTTGCTGCTCCGGCATCCGTTACAGAC
AAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTCACCGTCATACCGAAC
GCGCGA

FIGURE 99D

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

REC'D

A. The indications made below relate to the microorganism referred to in the description on page 54, line 8

B. IDENTIFICATION OF DEPOSITFurther deposits are identified on an additional sheet

Name of depositary institution

Agricultural Research Culture Collection (NRRL)
 International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street
 Peoria, Illinois 61604
 United States of America

Date of deposit February 27, 1999

Accession Number

NRRL B-30103

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*)This information is continued on an additional sheet

Escherichia coli DB3.1(pEZC15101)

In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

For International Bureau use only

 This sheet was received with the international application

 This sheet was received by the International Bureau on:

Authorized officer



Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 55, line 16.

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution

Agricultural Research Culture Collection (NRRL)
 International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street
 Peoria, Illinois 61604
 United States of America

Date of deposit February 27, 1999

Accession Number

NRRL B-30100

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*)

This information is continued on an additional sheet

Escherichia coli DB3.1(pENTR-1A)

In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*If the indications are not for all designated States*)**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the international Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

For International Bureau use only

This sheet was received with the international application

This sheet was received by the International Bureau on:

Authorized officer

Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

<p>A. The indications made below relate to the microorganism referred to in the description on page <u>55</u>, line <u>16</u>.</p>		
<p>B. IDENTIFICATION OF DEPOSIT</p>		<p>Further deposits are identified on an additional sheet <input checked="" type="checkbox"/></p>
<p>Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority</p>		
<p>Address of depositary institution (<i>including postal code and country</i>) 1815 N. University Street Peoria, Illinois 61604 United States of America</p>		
Date of deposit	February 27, 1999	Accession Number NRRL B-30102
<p>C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)</p>		<p>This information is continued on an additional sheet <input type="checkbox"/></p>
<p>Escherichia coli DB3.1(pENTR-3C)</p>		
<p>In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).</p>		
<p>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)</p>		
<p> </p>		
<p>E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>)</p>		
<p>The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i>)</p>		

<p>For receiving Office use only</p>		<p>For International Bureau use only</p>	
<p><input checked="" type="checkbox"/> This sheet was received with the international application</p>		<p><input type="checkbox"/> This sheet was received by the International Bureau on:</p>	
<p>Authorized officer <i>B. Indie</i></p>		<p>Authorized officer</p>	

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 55, line 16.

B. IDENTIFICATION OF DEPOSITFurther deposits are identified on an additional sheet

Name of depositary institution

Agricultural Research Culture Collection (NRRL)
 International Depository Authority

Address of depositary institution (*including postal code and country*)

1815 N. University Street
 Peoria, Illinois 61604
 United States of America

Date of deposit February 27, 1999Accession Number NRRL B-30101**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*)This information is continued on an additional sheet

Escherichia coli DB3.1(pENTR-2B)

In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the international Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

For International Bureau use only

 This sheet was received with the international application This sheet was received by the International Bureau on:

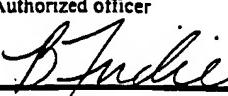
Authorized officer



Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>WPO</u> , <u>PCT</u> <u>20-21</u>	
B. IDENTIFICATION OF DEPOSIT <div style="float: right;"><input checked="" type="checkbox"/> Further deposits are identified on an additional sheet</div> <p>Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority</p> <p>Address of depositary institution (<i>including postal code and country</i>) 1815 N. University Street Peoria, Illinois 61604 United States of America</p>	
Date of deposit	February 27, 1999
	Accession Number NRRL B-30108
C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>) <div style="float: right;"><input type="checkbox"/> This information is continued on an additional sheet</div> <p>Escherichia coli DB10B(pCMVSPORT6).</p> <p>In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>) <div style="height: 150px; width: 100%;"></div>	
E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>) <div style="height: 150px; width: 100%;"></div> <p>The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i>)</p>	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer 	
<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

<p>A. The indications made below relate to the microorganism referred to in the description on page <u>54</u>, line <u>9</u>.</p>		
<p>B. IDENTIFICATION OF DEPOSIT</p>		<p>Further deposits are identified on an additional sheet <input checked="" type="checkbox"/></p>
<p>Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority</p>		
<p>Address of depositary institution (<i>including postal code and country</i>) 1815 N. University Street Peoria, Illinois 61604 United States of America</p>		
Date of deposit	February 27, 1999	Accession Number
		NRRL B-30105.
<p>C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)</p>		<p>This information is continued on an additional sheet <input type="checkbox"/></p>
<p>Escherichia coli DB3.1(pEYC15103)</p>		
<p>In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).</p>		
<p>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>If the indications are not for all designated States</i>)</p>		
<p>E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>)</p>		
<p>The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i>)</p>		

<p>For receiving Office use only</p>		<p>For International Bureau use only</p>	
<p><input checked="" type="checkbox"/> This sheet was received with the international application</p>		<p><input type="checkbox"/> This sheet was received by the International Bureau on:</p>	
<p>Authorized officer</p> 		<p>Authorized officer</p>	

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

<p>A. The indications made below relate to the microorganism referred to in the description on page <u>54</u>, line <u>9</u>.</p>		
<p>B. IDENTIFICATION OF DEPOSIT</p>		<small>Further deposits are identified on an additional sheet <input checked="" type="checkbox"/></small>
<p>Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority</p>		
<p>Address of depositary institution (<i>including postal code and country</i>) 1815 N. University Street Peoria, Illinois 61604 United States of America</p>		
Date of deposit	February 27, 1999	Accession Number
		NRRL B-30104.
<p>C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>)</p>		<small>This information is continued on an additional sheet <input type="checkbox"/></small>
<p>Escherichia coli DB3.1(pEZA15102)</p>		
<p>In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).</p>		
<p>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>if the indications are not for all designated States</i>)</p>		
<p>E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>)</p>		
<p>The indications listed below will be submitted to the International Bureau later (<i>specify the general nature of the indications, e.g.: "Accession Number of Deposit"</i>)</p>		

<p>For receiving Office use only</p>		<p>For International Bureau use only</p>	
<p><input checked="" type="checkbox"/> This sheet was received with the international application</p>		<p><input type="checkbox"/> This sheet was received by the International Bureau on:</p>	
<p>Authorized officer</p>		<p>Authorized officer</p>	
			

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

<p>A. The indications made below relate to the microorganism referred to in the description on page <u>52</u>, line <u>31</u>.</p>		
<p>B. IDENTIFICATION OF DEPOSIT</p>		<input checked="" type="checkbox"/> Further deposits are identified on an additional sheet
<p>Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority</p>		
<p>Address of depositary institution (<i>including postal code and country</i>) 1815 N. University Street Peoria, Illinois 61604 United States of America</p>		
Date of deposit	February 27, 1999	Accession Number
<p>C. ADDITIONAL INDICATIONS (<i>leave blank if not applicable</i>) This information is continued on an additional sheet <input type="checkbox"/></p> <p>Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)</p> <p>In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).</p>		
<p>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (<i>If the indications are not for all designated States</i>)</p>		
<p>E. SEPARATE FURNISHING OF INDICATIONS (<i>leave blank if not applicable</i>)</p> <p>The indications listed below will be submitted to the international Bureau later (<i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i>)</p>		
<p>For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer <u>Barbara Fridie</u> PCT Operations - I/PD Team 1 703) 305-3747 (703) 305-3230 (FAX)</p>		
<p>For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>		

*Escherichia coli DB3.1(pENTR-3C)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pENTR-3C)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pENTR-2B)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pENTR-2B)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pENTR-2B)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pENTR-1A)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pENTR-1A)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pENTR-1A)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB10B(pCMVSport6)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB10B(pCMVSport6)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB10B(pCMVSPORT6)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pEZC15103)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pEZC15103)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pEZC15103)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pEZC15102)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pEZC15102)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pEZC15102)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pEZC15101)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

*Escherichia coli DB3.1(pEZC15101)***ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pEZC15101)***SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

*Escherichia coli DB3.1(pENTR-3C)***AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/05432

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :Please See Extra Sheet.
 US CL :435/912, 252.3, 320.1; 530/350; 536/ 23.1, 24.1
 According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/912, 252.3, 320.1; 530/350; 536/ 23.1, 24.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P ---	US 5,888,732 A (HARTLEY et al.) 30 March 1999, see entire document.	1-21, 25-30 36-38 -----
Y,P		22-24, 31-35
X - Y	HASAN et al. Escherichia coli genome targeting, I. Cre-lox-mediated in vitro generation of ori- plasmids and their in vivo chromosomal integration and retrieval. Gene. 1994, Vol. 150, pages 51-56, see entire document.	1-5, 10, 11, 19-21 ----- 15-18, 22-38
X - Y	KATZ et al. Site-specific recombination in Escherichia coli between the att sites of plasmid pSE211 from Saccharopolyspora erythraea. Mol. Gen. Genet. 1991, Vol. 227, pages 155-159, see entire document.	1-11, 19-21 ----- 15-18, 22-38

Further documents are listed in the continuation of Box C. See patent family annex.

A	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"P"	document referring to an oral disclosure, use, exhibition or other means		
	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
08 MAY 2000	23 MAY 2000

Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer <i>Doretha Lawrence Fox</i> IREM YUCEL Telephone No. (703) 308-0196
---	---

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/05432

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ASTUMIAN et al. Site-specific recombination between cloned attP and attB sites from the Haemophilus influenzae bacteriophage HP1 propagated in recombination deficient Escherichia coli. J of Bacteriology. March 1989, Vol. 171, No. 3, pages 1747-1750, see entire document.	1-11, 19-21 -----
-		
Y		15-18, 22-38

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/05432

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (7):

C07H 21/04; C07K 1/00, 14/00; C12N 1/21, 15/00, 15/09, 15/63, 15/70; C12P 19/34

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

WEST, STN (CAPLUS); DIALOG (MEDLINE, BIOSIS, SCISEARCH, PASCAL)

Terms: att (B?, P?, R?, L?), MCS, POLYLINKER, PLASMID, VECTOR, LOCALIZATION, SIGNAL, TRANSCRIPTION, TERMIN?, TRANSLATION?, ORI, REPLICON, GST, HEXHIST?, THIOREDOX?, CLEAVAGE, SITE?, SPECIF?, DIRECT?, RECOMBIN?, CLON?, INSERT?

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.